



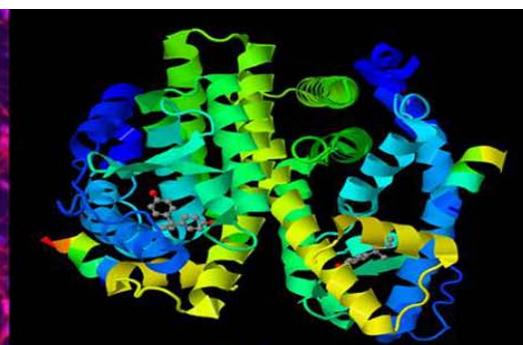
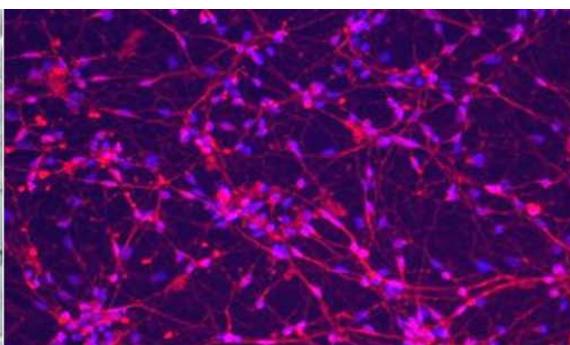
European
Commission

JRC SCIENCE AND POLICY REPORTS

Alternative methods for regulatory toxicology – a state-of-the-art review

A. Worth, J. Barroso, S. Bremer, J. Burton, S. Casati, S. Coecke, R. Corvi, B. Desprez, C. Dumont, V. Gouliarmou, M. Goumenou, R. Gräpel, C. Griesinger, M. Halder, A. Janusch Roi, A. Kienzler, F. Madia, S. Munn, M. Nepelska, A. Paini, A. Price, P. Prieto, A. Rolaki, M. Schäffer, J. Triebe, M. Whelan, C. Wittwehr, V. Zuang

2014



Joint
Research
Centre

European Commission

Joint Research Centre
Institute for Health and Consumer Protection
European Union Reference Laboratory for Alternatives to Animal Testing

Contact information

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 126, 21027 Ispra (VA), Italy
E-mail: JRC-ECVAM-CONTACT@ec.europa.eu

http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam

<https://ec.europa.eu/jrc>

Legal Notice

This publication is a Science and Policy Report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

All images © European Union 2014, unless otherwise indicated.

JRC91361

EUR 26797 EN

ISBN 978-92-79-39651-9 (PDF)

ISSN 1831-9424 (online)

doi:10.2788/11111

Luxembourg: Publications Office of the European Union, 2014

© European Union, 2014

Reproduction is authorised provided the source is acknowledged.

Abstract

This state-of-the art review is based on the final report of a project carried out by the European Commission's Joint Research Centre (JRC) for the European Chemicals Agency (ECHA). The aim of the project was to review the state of the science of non-standard methods that are available for assessing the toxicological and ecotoxicological properties of chemicals. Non-standard methods refer to alternatives to animal experiments, such as in vitro tests and computational models, as well as animal methods that are not covered by current regulatory guidelines.

This report therefore reviews the current scientific status of non-standard methods for a range of human health and ecotoxicological endpoints, and provides a commentary on the mechanistic basis and regulatory applicability of these methods. For completeness, and to provide context, currently accepted (standard) methods are also summarised. In particular, the following human health endpoints are covered: a) skin irritation and corrosion; b) serious eye damage and eye irritation; c) skin sensitisation; d) acute systemic toxicity; e) repeat dose toxicity; f) genotoxicity and mutagenicity; g) carcinogenicity; h) reproductive toxicity (including effects on development and fertility); i) endocrine disruption relevant to human health; and j) toxicokinetics. In relation to ecotoxicological endpoints, the report focuses on non-standard methods for acute and chronic fish toxicity. While specific reference is made to the information needs of REACH, the Biocidal Products Regulation and the Classification, Labelling and Packaging Regulation, this review is also expected to be informative in relation to the possible use of alternative and non-standard methods in other sectors, such as cosmetics and plant protection products.



EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)

Alternative methods for regulatory toxicology – a state-of-the-art review

A. Worth, J. Barroso, S. Bremer, J. Burton, S. Casati, S. Coecke, R. Corvi, B. Desprez, C. Dumont, V. Gouliarmou, M. Goumenou, R. Gräpel, C. Griesinger, M. Halder, A. Janusch Roi, A. Kienzler, F. Madia, S. Munn, M. Nepelska, A. Painsi, A. Price, P. Prieto, A. Rolaki, M. Schäffer, J. Triebe, M. Whelan, C. Wittwehr, V. Zuang

European Union Reference Laboratory for Alternatives to Animal Testing
(EURL ECVAM),
Systems Toxicology Unit,
Institute for Health and Consumer Protection,
European Commission Joint Research Centre,
Ispra, Italy.

EXECUTIVE SUMMARY

This state-of-the art review is based on the final report of a project carried out by the European Commission's Joint Research Centre (JRC) for the European Chemicals Agency (ECHA) under the terms of a Service Level Agreement (SLA/ECHA-JRC/2012-2) between ECHA and the JRC. The aim of the project was to review the state of the science of non-standard methods that are available for assessing the toxicological and ecotoxicological properties of chemicals. Non-standard methods refer to alternatives to animal experiments, such as *in vitro* tests and computational models, as well as animal methods that are not covered by current regulatory guidelines.

ECHA needs to have up-to-date information on non-standard methods since the Agency is responsible for implementing different regulatory processes in which there is an obligation or an opportunity to use non-standard methods, depending on the context. These processes relate to the REACH Regulation, the Biocidal Products Regulation (BPR), as well as the Classification, Labelling and Packaging (CLP) Regulation.

This report reviews the current scientific status of non-standard methods for a range of human health and ecotoxicological endpoints, and provides a commentary on the mechanistic basis and regulatory applicability of these methods. For completeness, and to provide context, currently accepted (standard) methods are also summarised. In particular, the following human health endpoints are covered: a) skin irritation and corrosion; b) serious eye damage and eye irritation; c) skin sensitisation; d) acute systemic toxicity; e) repeat dose toxicity; f) genotoxicity and mutagenicity; g) carcinogenicity; h) reproductive toxicity (including effects on development and fertility); i) endocrine disruption relevant to human health; and j) toxicokinetics. In relation to ecotoxicological endpoints, the report focuses on non-standard methods for acute and chronic fish toxicity.

While this report makes specific reference to the information needs of REACH, the BPR, and the CLP Regulation, it is also expected to be informative in relation to the possible use of alternative and non-standard methods in other sectors, such as cosmetics and plant protection products.

DISCLAIMER

This report presents the findings and conclusions of the authors, but does not represent an official view of the JRC, the European Commission or the European Chemicals Agency.

CONTENTS

1. Background to the report	1
2. Adverse Outcome Pathways and their role in assessing non-standard methods	3
3. Skin corrosion and irritation	7
4. Serious eye damage and eye irritation	59
5. Skin sensitisation	93
6. Acute systemic toxicity	116
7. Repeat dose systemic toxicity	142
8. Genotoxicity and mutagenicity	178
9. Carcinogenicity	211
10. Reproductive toxicity – effects on fertility and developmental toxicity	264
11. Endocrine disruption relevant to human health and wildlife populations	310
12. Toxicokinetics	334
13. Aquatic (fish) toxicity	368
Appendix 1 Weblinks to general information on non-standard methods	443
Appendix 2 Status of test method submissions to EURL ECVAM	445
Appendix 3 Status of computational models in the JRC QSAR database	455
Appendix 4 Status of alternative methods in DB-ALM	457

1. Background to the report

This state-of-the-art review is based on the final report of a project carried out by the European Commission's Joint Research Centre (JRC) for the European Chemicals Agency (ECHA) under the terms of a Service Level Agreement (SLA/ECHA-JRC/2012-2) between ECHA and the JRC. The aim of the project was to review the state of the science of non-standard methods that are available for assessing the toxicological and ecotoxicological properties of chemicals. Non-standard methods refer to alternatives to animal experiments, such as *in vitro* tests and computational models, as well as animal methods that are not covered by current regulatory guidelines. This project was one of several activities carried out within the frame of the collaboration agreement between the JRC's Institute for Health & Consumer Protection (IHCP) and ECHA.

ECHA needs to have up-to-date information on non-standard methods since the Agency is responsible for implementing different regulatory processes in which there is an obligation or an opportunity to use non-standard methods, depending on the context. These processes relate to the REACH Regulation, the Biocidal Products Regulation (BPR), and well as the Classification and Labelling and Packaging (CLP) Regulation.

The registration of chemicals under REACH (Regulation 1907/2006 on the Registration, Evaluation and Authorisation of Chemicals) is based on their production volume and their hazard properties (i.e. Substances of Very High Concern [SVHCs]), with the assessment of the inherent substance properties based on data provided by manufacturers or importers. The registrant is encouraged to provide data obtained with non-standard methods under the conditions described in Annex XI of the REACH legal text. When this is done it is necessary to demonstrate that the data provided are relevant, reliable, and that their use in the context of classification and risk assessment is acceptable. REACH registration dossiers may be examined by ECHA under the dossier evaluation processes.

Under CLP (Regulation 1272/2008 on the Classification, Labelling and Packaging of substances and mixtures), hazard information on chemicals is assessed against the classification criteria in order to establish a harmonised classification which is legally binding within the EU, or to provide a self-classification for hazard classes or categories not falling under the harmonised classification procedure. Self-classifications are reported by manufacturers and importers to the CLP Inventory, hosted by ECHA, for all chemicals placed on the market or subject to REACH, however they are not legally binding. Under CLP, it is possible to use non-standard data as a basis for hazard assessment in a weight of evidence approach.

The use of non-standard methods is also pertinent to hazard and risk assessments carried out under the BPR (Regulation (EU) 528/2012), which concerns the placing on the market and use of biocidal products.

Furthermore, in the light of the Directive on the protection of animals used for scientific purposes (Directive 2010/63/EU), the use of non-standard methods not requiring the use of animals is encouraged in all sectors of EU Chemicals Policy. It is however fundamentally important to ensure that the use of non-standard data is appropriate for an adequate risk assessment or classification and does not reduce the protection of human health and the environment.

This report reviews the current scientific status of non-standard methods for a range of human health and ecotoxicological endpoints, and provides a commentary on the mechanistic basis and regulatory applicability of these methods. For completeness, and to provide context, currently accepted (standard) methods are also summarised. In particular, the following human health endpoints are covered: a) skin irritation and corrosion; b) serious eye damage and eye irritation; c) skin sensitisation; d) acute systemic toxicity; e) repeat dose toxicity; f) genotoxicity and mutagenicity; g) carcinogenicity; h) reproductive toxicity (including effects on development and fertility); i) endocrine disruption relevant to human health; and j) toxicokinetics. In relation to ecotoxicological endpoints, the report focuses on non-standard methods for acute and chronic fish toxicity.

While this report makes specific reference to the information needs of REACH, the BPR, and the CLP Regulation, it is also expected to be informative in relation to the possible use of alternative and non-standard methods in other sectors, such as cosmetics and plant protection products.

2. Adverse Outcome Pathways and their role in assessing non-standard methods

Andrew Worth, Sharon Munn, Maurice Whelan & Clemens Wittwehr

2.1 The Adverse Outcome Pathway (AOP) concept

Over the past decade, attempts to describe and predict the biological and toxicological effects of chemicals have increasingly taken mechanistic considerations into account. Among the many publications in the rapidly growing field of predictive toxicology, the report of the U.S. National Academy of Sciences and National Research Council (NRC) “Toxicity Testing in the 21st Century: A Vision and A Strategy” (NRC, 2007) has attracted considerable attention and triggered a debate about the need for a “paradigm shift” in hazard and risk assessment of chemicals (Collins et al, 2008; Andersen et al, 2010; Ankley et al, 2010; Schultz, 2010; Hartung & Bride, 2011; Boekelheide & Andersen, 2010). The NRC report envisions a transformation of the current way of conducting toxicity testing from a system based on phenotypic responses in animals towards an approach that increasingly relies on an understanding of the molecular mechanisms of toxicant effects in human cells and tissues. The move towards a more mechanistically-based risk assessment process implies the use of in vitro tests based on high-throughput and high-content screening (HTS/HCS) assays in mammalian (preferably human) cell lines, cell cultures and/or tissue surrogates, combined with the application of a range of computational methods for data analysis and the modelling of molecular interactions between toxicant and molecular target, adverse effects and fate.

Different terms are being used to capture variants of this general framework, including source-to-outcome pathway, toxicity pathway (TP), mode of action (MoA) and adverse outcome pathway (AOP). While these terms are not yet harmonised, they are all based on the assumption that a toxicant, after reaching and interacting with a biological target, initiates a cascade of events which may lead to an adverse outcome at the organism or population level. The general premise of the AOP approach is that a limited set of key measurable events are sufficient for describing biological pathways and predicting adverse outcomes at multiple levels of biological organisation (cell, tissue/organ, organism, population). For practical purposes in chemical hazard and risk assessment, this means that a basic understanding of the key molecular interactions and effects should be sufficient, and that ultimately it may be sufficient for decision making to predict the adverse outcome at organism and population level from early (“upstream”) key events. Table 2.1 provides definitions for some commonly used terms, and also includes working definitions for the purposes of this report.

Table 2.1. Definitions used in mechanistic frameworks for predicting and assessing toxicity

Term	Definition	Reference
Key Event	Key events are events that are toxicologically relevant and necessary to the apical outcome and experimentally quantifiable. They include the Molecular Initiating Event and subsequent intermediate effects.	Definition for the purpose of this report
Molecular Initiating Event (MIE)	The initial interaction of a chemical and a biological target which results in a cascade of events (perturbation of the biological system) and which may lead to adversity at a higher level of biological organisation.	Definition for the purpose of this report
Intermediate Effect	An event in a pathway that occurs after (downstream of) the Molecular Initiating Event and is part of the cascade which may lead to adversity at a higher level of biological organisation. An intermediate effect is distinguished from a key event in that not all intermediate effects are necessarily key events.	Definition for the purpose of this report
Source-to-Outcome Pathway	Cascade of measurable events starting from release of a chemical into the environment to an adverse outcome.	Krewski et al, 2011
Toxicity pathway (TP)	A cellular response pathway that would result in an adverse health effect when sufficiently perturbed	NRC, 2007
Mode of Action (MoA)	Sequence of events, starting with a Molecular Initiating Event, and leading to an adverse effect at the level of whole organism/individual. This term does not (usually) include consideration of exposure or effects at higher levels than the individual.	EPA, 2005
Mode of Action (MoA)	A biologically plausible sequence of key events leading to an observed effect supported by robust experimental observations and mechanistic data. A mode of action describes key cytological and biochemical events – that is, those that are both measurable and necessary to the observed effect – in a logical framework.	WHO, 2009
Adverse Outcome Pathway (AOP)	The sequence of events from the exposure of an individual or population to a chemical substance through a final adverse (toxic) effect at the individual level (for human health) or population level (for ecotoxicological endpoints). AOPs incorporate the toxicity pathway and mode of action for an adverse effect.	OECD, 2013

2.2 Use of the AOP framework in assessing non-standard methods

The AOP framework provides a means of organising and sharing knowledge on the mechanisms underlying chemical toxicity. It is a descriptive framework that can be used as the basis of predictive approaches, including chemical categories, mechanistically based QSARs, *in vitro* tests with toxicologically relevant read-outs, as well as integrated approaches (often referred to as Integrated Testing Strategies [ITS] or Integrated Approaches to Testing and Assessment [IATA]). A recently published OECD Guidance Document (OECD, 2013) provides guidance on which pieces of information are necessary to identify, document and justify an AOP in terms of its relevance and adequacy. A reporting template is also provided to improve consistency in the reporting of AOPs.

In this report, we will use the elements of the AOP framework, as far as is possible in the light of incomplete knowledge and the absence of an agreed ontology of key events, to identify non-standard methods that are toxicologically relevant. In other words, these are non-standard methods that can be associated with key events in one or more AOPs.

2.3 Development of an AOP Knowledge Base

Progress made with the implementation of the AOP framework will generate data and knowledge describing the molecular initiating events (MIEs), Intermediate Effects (IEs) and Adverse Outcomes (AOs). In order to facilitate the collection and retrieval of this information in a coherent way, the JRC has instigated and is involved in a series of initiatives that will help generating a flexible yet standardised way of looking at AOP related data.

Being committed to the standard data format paradigm of OHTs (OECD Harmonised Templates, a collection of data entry forms for 100+ endpoints describing a chemical), the JRC has developed an OHT (working title "OHT 201", assigned by OECD) which will, once adopted by OECD, allow capturing chemical-specific data on IEs, to be used for hazard assessment in all kinds of contexts. In addition, OHT 201 formatted study reports of effects can be used as real-life manifestations of Key Events playing roles in one or more AOPs, i.e. not necessarily linked to one single Adverse Outcome (as is the case with the current OHTs).

Consequently, Intermediate Effects reported using OHT 201 will be linked together to form AOPs. At the request of the OECD EAG MST (Extended Advisory Group on Molecular Screening and Toxicogenomics) and WHO IPCS (World Health Organisation International Programme on Chemical Safety), the JRC is implementing (together with US EPA) an AOP Knowledge Base (AOP-KB), which will consist of three elements:

- AOP Wiki, a text-based tool implementing the OECD AOP guidance, to be used for capturing qualitative AOPs, being the first module to be exposed to the public;
- AOP Effectopedia, a graphical tool for capturing quantitative AOPs (but synchronized with the qualitative AOP wiki data) and foreseen to depict mathematical AOP models;
- AOP Intermediate Effects Database - a collection of IEs captured in the OHT 201 format.

2.4 References

Andersen ME, Al-Zoughool M, Croteau M, Westphal M & Krewski D (2010). The Future of Toxicity Testing. *Journal of Toxicology and Environmental Health, Part B* 13, 163-196.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tietge JE & Villeneuve DL (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 29, 730-741.

Boekelheide K & Andersen ME (2010). A mechanistic redefinition of adverse effects - a key step in the toxicity testing paradigm shift. *ALTEX* 27, 243-52.

Collins FS, Gray GM & Bucher JR (2008). Toxicology. Transforming environmental health protection. *Science* 319, 906-907.

Hartung T & McBride M (2011). Food for Thought ... on mapping the human toxome. *ALTEX* 28, 83-93.

Krewski D, Westphal M, Al-Zoughool M, Croteau MC & Andersen ME (2011). New Directions in Toxicity Testing. *Annual Review of Public Health* 32, 161-78.

NRC (2007). Toxicity Testing in the 21st Century: A Vision and A Strategy. 2007, National Academic Press: Washinton, DC.

OECD (2013). Guidance Document On Developing and Assessing Adverse Outcome Pathways. OECD Environment, Health and Safety Publications. Series on Testing and Assessment No. 184. ENV/JM/MONO(2013)6. Available at: <http://www.oecd.org/chemicalsafety/>

Schultz TW (2010). Adverse outcome pathways: a way of linking chemical structure to in vivo toxicological hazards. In *In Silico Toxicology: Principles and Applications* (ed. M.T.D. Cronin & J.C. Madden), pp. 346-371. Cambridge, UK: The Royal Society of Chemistry.

US EPA (2005). Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001F March 2005. Washington, DC: US Environmental Protection Agency. Available from: http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF

WHO (2009). Environmental Health Criteria 240: Principles and Methods for the Risk Assessment of Chemicals in Food. World Health Organization, Geneva. Available from: <http://www.who.int/entity/foodsafety/chem/principles/en/index.html>

3. Skin corrosion and irritation

Claudius Griesinger, Michael Schaeffer, Andrew Worth & Valérie Zuang

3.1 Introduction

Being the primary interface between the human body and the environment, the skin is readily exposed to chemicals. In the industrialised world, adverse skin effects are of major concern for occupational and consumer safety (Elsner, 1994; Wigger-Alberti & Elsner, 2000). Irritant Contact Dermatitis (ICD), resulting from acute or cumulative skin irritation is, after skeletomuscular diseases, the most common occupational disease. There is reliable evidence that ICD has, at least in the workplace, a greater incidence (Turner et al., 2007; Dickel et al., 2002) than Allergic Contact Dermatitis (ACD) resulting from skin sensitisation. Consequently, legislators and regulators attach high priority to the identification of chemicals that have skin corrosion or irritation potential so that adequate risk management measures can be chosen to protect workers and consumers. Despite the sophisticated barrier properties of skin, a wide range of chemicals that skin is acutely or repeatedly exposed to can lead to adverse effects in the skin tissue. This is reflected in the information requirements of the REACH Regulation, where data on skin irritation/corrosion is required for registrations at the lowest tonnage level of 1-10 tonnes per year. According to the severity and reversibility of effects one distinguishes skin corrosion (skin burns) from skin irritation.

Skin corrosive substances damage the exposed skin area leading to necrosis (death) of the skin tissue beyond repair. As a consequence, the affected area can be regenerated only from the healthy skin surrounding the necrotic patch. For the purposes of chemical safety assessments, skin corrosion is defined on the basis of the traditional animal test (not on the basis of effects in humans): *'Dermal [skin] corrosion is the production of irreversible damage of the skin; namely visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discolouration due to blanching of the skin, complete areas of alopecia and scars'*. (Citation from TG404; the same definition is reproduced in the EU CLP Regulation [EC 2008b] based on the United Nations UN Globally Harmonised System [UN GHS; UN 2013]).

Skin irritation is due to a reversible local inflammatory reaction in the skin. In contrast to corrosion, it involves complex and yet not fully described interactions of the innate immune system of the affected tissue patch. The clinical term for skin irritation, as observed in humans, is Irritant Contact Dermatitis (ICD). The inflammatory response leads, as downstream effects, to the classical clinical symptoms of irritation: redness (rubor), warmth (calor), painful sensations (dolor) and also the swelling (oedema) of the affected tissue area. For the purposes of chemical safety assessments, skin irritation is defined on the basis of the traditional animal test (not on the basis of the effects in humans): *'Dermal [skin] irritation is the production of reversible damage of the skin following the application of a test substance for up to 4 hours.'* (Citation from TG404; the same definition is reproduced in the EU CLP Regulation [EC 2008b] based on UN GHS [UN 2013]).

While some (mainly intrinsically weak irritant) chemicals will only trigger an irritant response after repeated exposure of the same skin area (cumulative irritants) (Dahl

1988), other chemicals will even after a one-time exposure cause irritation (acute irritants). Chronic cumulative exposure to irritants can elicit dermatitis which may resemble Allergic Contact Dermatitis and lead to inappropriate treatment and/or preventive measures (Dahl 1988; Basketter et al., 2004b). Although cumulative irritation clinically is far more frequent than acute irritation, current regulatory requirements focus on the assessment of the acute irritation potential of chemicals (i.e. intrinsic hazard of causing skin irritation following acute exposure). Therefore, this report focuses on acute irritation.

Information on skin irritation may be relevant for other health effects under certain circumstances. There is indication from human studies that the early inflammatory response of skin to sensitising and irritant substances is the same (Willis et al., 1986), and dermal inflammation is one of the key events involved in the skin sensitisation pathway (OECD, 2012). Thus, it is conceivable that substances that cause irritation without being sensitisers themselves may create an inflammatory environment supporting the skin sensitisation process caused by other chemicals. Therefore, consideration of the irritant potential of substances contained in mixtures (e.g. cosmetic products) that also contain substances with sensitising properties is important.

3.2 The traditional animal test and its regulatory use

Both skin corrosion and irritation have been traditionally assessed using rabbits as models. The traditional *in vivo* assay, the Draize test (Draize et al., 1944) has been in use for well over half a century. OECD Test Guideline (TG) 404 (OECD 2002) summarises the internationally recognized standard procedure for the testing of the dermal toxicity potential of substances *in vivo*, covering both corrosives and irritants. Several improvements with regard to its protocol and the conclusions drawn have been introduced over time (adoption as OECD TG in 1981 with revisions in 1992 and 2002, when a testing strategy supplement was added). Using the live animal, the effects of substances on patches of shaved skin that have been exposed for up to four hours are assessed and scored for (a) erythema (skin reddening) and eschar formation as well as (b) oedema (skin swelling). The scoring uses five steps, from 0 (no erythema / oedema) to four (maximum effect). Corrosion is differentiated from irritation by virtue of non-reversibility of the effects and by the production of serious histopathological damage to the skin in case of corrosion (ulcers, bleeding, discolouration, blanching of skin, alopecia, scars) at the end of the observation period (14 days).

The Draize method has the advantage of providing a simple readout of corrosive effects as well as the downstream effects of the inflammatory response induced by irritant xenobiotics (i.e. reddening, swelling etc). The inflammatory response ultimately leads to observable phenomena, two of which are assessed in the traditional animal test: localised skin swelling (edema - due to increased permeability of blood vessels to facilitate diapedesis/extravasation of immune cells into the interstitium) and redness (erythema - due to the increased diameter of blood vessels). Other reactions (e.g. C fibre activation, pain) are not easily accessible through such an observational method.

Despite these advantages, the *in vivo* method also has methodological drawbacks, the most important ones relating to variability and relevance regarding the species of interest: (a) coloured substances are difficult to score due the visual inspection of the skin and skin redness being a readout; (b) being performed in a proxy model (the rabbit) the test may make incorrect predictions due to species differences (e.g. Philips et al.

1972; Basketter et al. 2004b, Jirova et al., 2010; Basketter et al., 2012); (c) probably due to the fact that the test relies on the subjective scoring of the two effects (instead of using an empirical measure of a parameter related to irritation) variability of the recorded responses is high (e.g. Weil & Scala, 1971; ECETOC, 1995); this variability may be exacerbated due to inter-individual (inter-animal) variability with regard to the severity of the responses.

Data on skin irritation/corrosion effects are required by several pieces of legislation, all of which regulate use classes of chemicals that are likely to come into contact with skin: a) the REACH Regulation (EC 1907/2006); b) the Biocides Regulation (EC 528/2012); c) the Plant Protection Product Regulation (EC 1107/2009); and d) the Cosmetics Regulation (EC 1223/2009; in force since 11 July 2013, replacing Directive 76/768/EEC). Such legislation foresees the development and use of alternatives to traditional animal testing. The REACH regulation and accompanying ECHA guidance, for example, specifies the criteria by which registrants may be able to meet the information requirement for an animal test for skin irritation by using results obtained from human experience and alternative approaches where these are available. It may also be possible to combine the use of such information in a weight of evidence approach.

The CLP Regulation (Regulation on Classification, Labelling and Packaging of substances and mixtures: EC 1272/2008) does not itself stipulate any information requirements, but it lays down the rules for classification and labelling of skin corrosion and irritation that are applicable to REACH. The CLP Regulation implements the UN Globally Harmonised System (UN 2013) for C&L and repeals Directives 67/548/EEC ("Dangerous Substances Directive") and 1999/45/EC ("Dangerous Preparations Directive") and provides the classification criteria as well as the categories and subcategories for skin corrosion and irritation. Currently the CLP Regulation foresees Categories 1A, 1B and 1C for corrosive substances and Category 2 for severe irritants. Category 3 (mild irritants) is not implemented in the EU. Although EU CLP describes the general Category 1 (section 3.2.2.6.1 and Table 3.2.1 therein), there is some inconsistency regarding the implementation of a general Category 1 as UN GHS originally foresaw its use for authorities not implementing subcategories ("applies to authorities not using sub categories", UN 2011). Recently the 5th revision of UN GHS has been published which allows the use of Category 1 also for countries that do implement subcategories (UN 2013). This change will likely be taken up in EU CLP.

More details on these information requirements as well as the relevant categories as stipulated in the CLP Regulation are given below (see References with Notes).

3.2.1 Variability of the Draize skin test

A systematic analysis of variability of Draize test data was performed in 1971 (Weil and Scala, 1971): 10 substances were distributed to 24 laboratories and intra- and inter-laboratory variability of the scoring was assessed. The study found moderate within laboratory reproducibility and low between laboratory reproducibility and concluded that some of the substantial variability observed may be due to (1) the subjective way of scoring effects and (2) variations between laboratories in performing the test. However, high variability is also evident in the ECETOC database of skin irritation chemicals. Since these data were all produced following OECD test guideline 404 and under Good Laboratory Practice (GLP), variations due to between-laboratory deviations in the

test protocols can be excluded with high certainty. Hence, the variability is most likely based on either a) the subjective scoring or b) the intrinsic variability of responses in animals or c) a combination of both factors.

One of the reasons for employing only one irritant category (category 2) under UN GHS (category 3 is an optional category) is inter-animal variability: the UN GHS text explicitly acknowledges that "...animal responses in a test may be quite variable" in the context of explaining the rationale for one single irritant category (UN, 2013).

Inter-animal (within test) and between test (laboratory) variability is also evident in the high quality dataset of the ECETOC skin irritation chemicals (ECETOC, 1995). Based on this dataset, Griesinger et al. (2010) show that the variability in the erythema and oedema scores can be observed across the entire range of Draize scores (from 0 to 4). These quality-assured data show high inter-animal variability, and it is not immediately evident how the traditional *in vivo* test can indeed be used for sub-categorisation beyond one irritant category. The variability in Draize skin scores is also analysed by Worth & Cronin (2001), who illustrate its impact on the maximum predictive performance that can be expected of any alternative method trying to reproduce such scores.

The variability of the animal test places an upper limit on the expected performance of alternative test methods that were mainly developed, optimised, evaluated and validated in reference to Draize data (Worth & Cronin, 2001a). As the majority of alternative *in vitro* methods are based on human keratinocytes, they may potentially resemble more closely the human situation. Recent studies indicate that the alternative methods for skin irritation testing based on Reconstructed human Epidermis (RhE) are more accurate with regard to predicting effects in humans than the traditional animal test (Jirova et al., 2012). ECVAM has taken this into account when defining the Performance Standards (ECVAM, 2009a) for OECD TG 439 (EU test method B.46), where some of the reference predictions are based on human data from the human patch test (Basketter et al., 2004b; Basketter et al., 2012).

3.3 Mechanistic understanding of the endpoints

3.3.1 Skin corrosion

Mechanisms of skin corrosion concern mainly direct destruction of the skin tissue by chemicals, i.e. the erosion of the stratum corneum by many inorganic acids and bases and by strong organic acids with extreme pH: pH < 2 and > 11.5. Chemical classes that are taken up by skin and react with proteins, lipids and other biomolecules of skin tissue include quaternary ammonium ions, heterocyclic ammonium ions, sulphonium salts and acrylates. Cationic surfactants are readily taken up by skin and may bind to skin components and disrupt cellular plasma membranes thus leading to erosion of the tissue and subsequent necrosis.

3.3.2 Skin irritation

Skin irritation triggered by chemicals was originally assumed to be a simple process but is now considered a complex biological event with distinct pathophysiology and involving a complex reaction of the skin's innate immune system that is far from being fully understood and comprehensively described. Comprehensive reviews on the pathophysiology of skin irritation have been published (Basketter et al., 1997; Wells et al., 2004; Chew & Maibach, 2006; Kindt et al., 2006; Fluhr et al., 2008).

A putative Adverse Outcome Pathway (AOP) for irritation is being developed by the JRC (Griesinger, in preparation) based on the following key events leading to chemically-induced skin inflammation: dermal bioavailability and damage to the dermal barrier, metabolism, chemically-induced tissue trauma, release of inflammatory mediators, activation of the innate immune system). This AOP is elaborated in Appendix 1.

Although dermal inflammation has been identified as a key event of the OECD's Adverse Outcome Pathway for skin sensitisation, this should not lead to the conclusion that dermal irritation is mechanistically fully understood. In particular, the complex interactions of inflammatory mediators with the skin tissue, and the role of axonal reflexes (neurogenic inflammation) in the pathogenesis/toxicogenesis of skin irritation, are not understood in detail. Recently, irritation has been classified into 10 types based on the morphological and clinical course of the inflammatory reaction (Sugibayashi et al., 2002) suggesting that distinct mechanisms or pathways may be involved. There are forms of irritation that lack the classical clinical signs of inflammation described by Rudolf Virchow (redness, swelling, heat, pain and dysfunction). Such purely sensory or "subjective" irritation is nevertheless an important entity in the full range of clinical irritation (Lammintausta et al., 1988). Mechanistic investigations of irritation have profited from advanced imaging methods allowing the optical investigation of inflamed tissue (Astner et al., 2006; Welzel et al., 2003).

3.4 Status of alternative methods

For evaluation of skin corrosion and irritation, several *in vitro* and *ex vivo* test methods are accepted in addition to the traditional *in vivo* test (Draize rabbit test). The methods are presented in more detail in the following two subsections. Table 3.1 provides an overview of currently accepted test methods, and also indicates those standardized testing methods that have recently (in 2013) been updated in view of adaptation to technical progress such as amendment of performance standards or inclusion of non-validation data on predictive capacity. Briefly, these updates concerned OECD test guidelines TG 430 (addition of Performance Standards), TG 431 (extension of applicability domain to include also partial information on subcategorisation of Category 1 into 1A and a combination of Categories 1B and 1C) and TG 439 (inclusion of additional test method validated in reference to the TG's Performance Standards).

Table 3.1. EU test methods and OECD test guidelines for skin corrosion and irritation testing

Generic description of test method	Standardised description of test method in EU Test Method Regulation	OECD Test guidelines (TGs)
<i>In vivo</i> skin corrosion and irritation		
<i>In vivo</i> Draize rabbit test for skin corrosion/irritation testing	B.4 Acute toxicity: dermal irritation/corrosion	TG 404 Acute Dermal Irritation/Corrosion
<i>In vitro</i> skin corrosion		
Transcutaneous electrical resistance test (TER)	B.40 <i>In vitro</i> skin corrosion: Transcutaneous electrical resistance test (TER)	TG 430 <i>In vitro</i> skin corrosion: Transcutaneous electrical resistance test (TER) > Updated in 2013
<i>In vitro</i> skin corrosion using RhE-based human skin models: <ul style="list-style-type: none"> • Episkin™ • Epiderm™ • SkinEthic™ • EpiCS™ (EST-1000) 	B.40bis <i>In vitro</i> skin corrosion: human skin model test	TG 431 <i>In vitro</i> skin corrosion: human skin model test > Updated in 2013
CORROSITEX®	No EU test method available	TG 435 <i>In Vitro</i> Membrane Barrier Test Method for Skin Corrosion
<i>In vitro</i> skin irritation		
<i>In vitro</i> skin irritation testing using RhE-based human skin models: <ul style="list-style-type: none"> • Episkin™ • Epiderm™ SIT (EPI-200) • SkinEthic™ RHE • LabCyte™ EPI-MODEL24 SIT 	B.46 <i>In Vitro</i> Skin Irritation: Reconstructed Human Epidermis (RHE) Model Test. B.46 has not yet been updated to include the LabCyte™ test method.	TG 439 <i>In Vitro</i> Skin Irritation: Reconstructed Human Epidermis Test Method > Updated in 2013

3.4.1 Standard non-animal test methods for corrosion

Internationally accepted test methods for skin corrosion testing are, apart from the traditional animal test (TG404):

- TG435: the Corrositex™ test method based on an artificial membrane barrier (hydrogenated collagen matrix)
- TG430: the Transcutaneous Electrical Resistance (TER) test method based on excised skin
- TG431: *in vitro* test methods based on Reconstructed human Epidermis (RhE) using normal human keratinocytes.

Corrositex

The Corrositex test method is a simple *in chemico* test method based on fact that corrosive substances either pass through or destroy the skin barrier (composed mainly of extracellular matrix and lipid membrane layers in the stratum corneum), penetrating into

the tissue where they lead to severe burns. The Corrositex method recapitulates the stratum corneum barrier of the skin using a hydrated collagen matrix supported by a filter membrane. Corrosive chemicals or mixtures pass through this biobarrier (by diffusion and/or destruction/erosion) and elicit a colour change in the underlying liquid "Chemical Detection System" (CDS), composed of water and pH indicator dyes. The assay has the limitation that many non-corrosive chemicals and mixtures and some corrosive chemicals and mixtures may not qualify for testing with this test method. For instance, aqueous substances with a pH range from 4.5 to 8.5 do typically not qualify for testing in the Corrositex assay. This excludes a lot of potential corrosives. The Corrositex method has undergone evaluation and validation by ECVAM (Fentem et al., 1998; ECVAM 2000b) and NICEATM/ICCVAM (ICCVAM, 1999). While the Corrositex test method is accepted by OECD, it has not been included in the EU test methods regulation (EU, 2008a). The test method allows subcategorisation of corrosives (Category 1) into subcategories 1A, 1B, 1C, implemented by EU CLP and based on the UN GHS (UN, 2013), but, as explained above, its chemical applicability is limited.

Transcutaneous Electrical Resistance (TER)

The TER test method is based on the testing of chemicals on excised skin discs from rats to identify corrosives. The test method is based on the observation that corrosive substances impair the barrier function of skin by reacting with extracellular matrix molecules and destroying/eroding this matrix leading to a reduction of the electrical resistance across the skin. The test method was validated by ECVAM on the basis of about 60 chemicals of a wide range of chemical classes and physical states (Barratt et al., 1998; Fentem et al., 1998), including liquids (aqueous or non-aqueous), semi solids, solids (soluble or insoluble in water) and waxes. While mixtures were not included in the validation study, there is no obvious *a priori* reason to exclude them from the applicability domain of the test. TG430 provides more information on this issue and complicating factors (e.g. mixtures cover a wide spectrum of chemical categories and composition and there is considerable lack of availability of information on testing of mixtures) and points to a proposed strategy for testing mixtures developed during a recent workshop on regulatory applicability of skin corrosion/irritation test methods (Eskes et al., 2012). The test method does not allow the subcategorisation of corrosives (Category 1) into subcategories 1A, 1B, 1C implemented by EU CLP and based on the UN GHS (UN, 2013).

Reconstructed human Epidermis models for skin corrosion testing (TG 431)

RhE models use normal (e.g. non-transformed) human epidermal keratinocytes that, during culturing, form a multi-layered epidermis including a stratum corneum at the top, functioning as a barrier. These tissue-engineered models closely resemble human epidermis from a histological and biochemical point of view (e.g. lipid profile). The test method is based on the capacity of skin corrosives to destroy tissue and thus impair the viability of skin cells. Using a simple prediction model, the skin corrosion potential is predicted on the basis of cell viability measurements (using MTT/formazan) following exposure of the artificial epidermis and a post incubation period. At present the following RhE-based methods are accepted for skin corrosion testing:

- EpiSkin™
- EpiDerm™ SCT
- SkinEthic™ RHE
- epiCS®

Following prevalidation (Botham et al., 1995), the EpiSkin™ test method was validated by ECVAM in 1998 (Barratt et al., 1998; Fentem et al., 1998; ECVAM 1998). The EpiDerm™ test method underwent external validation in 2000 (Liebsch et al., 2000; ECVAM, 2000a). The epiCS® (Hoffmann et al., 2005) and SkinEthic™ (Kandarova et al., 2006; Tornier et al., 2010) test methods have been validated based on Performance Standards (ECVAM 2006; ECVAM 2009c).

All test methods can be used to distinguish corrosive substances from substances not requiring classification (ECVAM 1998, 2000a, 2000b, 2006, 2009c). In addition, the EpiSkin was also found capable to distinguish the categories R35 (now Category 1) from R34 (a combination of Categories 2 and 3). Recently the test method developers of the EpiSkin™, EpiDerm™ and SkinEthic™ test methods have re-evaluated (under non-validation conditions) the capacity of the test methods to subcategorise corrosives. More details are described in the "References with Notes" section. The applicability domain of these methods is as described for the TER and the same considerations regarding the testing of mixtures apply.

3.4.2 Standard non-animal test methods for irritation

Currently, all internationally accepted test methods for skin irritation testing (apart from the traditional animal test) are *in vitro* test methods based on the Reconstructed Human Epidermis (RhE) technology validated by ECVAM (TG439 / EU test method B.46). The test methods are shown in Table 3.2, which also outlines the type of validation study and provides the relevant references. All these test methods are proprietary, i.e. they need to be purchased from the test method developers:

- EpiSkin™
- EpiDerm™ SIT (EPI-200)
- SkinEthic™ RHE
- LabCyte™ EPI-MODEL24 SIT

The EpiSkin™ and the original EpiDerm™ test methods underwent full validation, while the updated version of the EpiDerm™ test method as well as the SkinEthic™ and LabCyte™ test methods underwent validation on the basis of the Performance Standards defined by ECVAM (ECVAM 2007 and 2009; Griesinger et al., 2010) and annexed to TG 439. A further validation study on the epiCS® model (formerly known as EST-1000 and already accepted for skin corrosion assessment) is expected to be validated by 2014). Furthermore, an open-source RhE-based skin model ("OS-Rep") based on the reconstructed human epidermis model and protocol developed by Poumay et al., (2004) is currently being validated by industry and will be submitted to ECVAM for evaluation and potential ECVAM Scientific Advisory Committee (ESAC) review. All technical details for producing/reconstructing this model are freely available.

All accepted RhE models use normal (e.g. non-transformed) human epidermal keratinocytes that, during culturing, form a multi-layered epidermis including a stratum corneum at the top, functioning as a barrier. These tissue-engineered models closely resemble human epidermis from a histological and biochemical point of view (e.g. lipid profile). Clearly however, the RhE systems lack vascularisations, i.e. have no blood vessels and, as they are not full thickness models, also the dermal innervation. Thus, they cannot recapitulate the endpoints assessed (through observation) in the animal test, which are due to the physiology of the endothelial (vascular) system in skin: erythema

and oedema. The RhE methods mechanistically assess at present cell and tissue damage/trauma, the primary event in the skin inflammation cascade leading to dermal irritation. This is measured through viability of the cells in the tissue constructs (i.e. keratinocytes). There have been attempts to also use the release of inflammatory mediators as readouts for making predictions on the skin irritation potential of chemicals (e.g. Interleukin 1 alpha). However, so far the variability of these parameters was too high to be taken as a basis for routinely used standardised test systems (Griesinger & Zuang 2012, ECVAM, 2007a).

All of the RhE test methods are accepted for predicting the presence (Category 2) and absence (no classification) of skin irritation potential based on the UN GHS system (UN, 2013) as implemented in the EU through the CLP Regulation (EU 2008b). However, none of the methods is able to predict the optional UN GHS Category 3 ("mild irritants"). While this optional Category is not implemented in the EU (EU, 2008b), it is used in other areas of the world (e.g. USA) and the extent to which data generated with TG439 methods can be used therefore depends on the regulatory context. In the EU, however, the TG439 test methods are considered full replacements to the animal test as they address the information requirements stipulated in relevant legislation (EU, 2008b). There are only very few restrictions known at present with regard to the chemical applicability domain of these methods. On the basis of physicochemical properties the test methods are applicable to: a) solids (water soluble and water insoluble); b) liquids (aqueous and non-aqueous); c) semi-solids; and d) waxes.

The chemicals used for validation were all from the EU new chemicals database (Eskes et al., 2007, Spielmann et al., 2007), representing molecules of a complex and modern chemistry. This may create confidence that the test methods are applicable also to a wide range of (also complex) chemical structures. However, highly coloured chemicals may interfere with the colorimetric cell viability measurement of the test methods. Other restrictions are currently not known.

Table 3.2. *In vitro* test methods accepted by OECD and EU for skin irritation testing, all based on Reconstructed human Epidermis technology

1. Validated test methods used for development of the Performance Standards			
Nr.	Test method	Validation study type	References
1	EpiSkin™	Full prospective validation study (2005-2007). The test method components of this method were used to define the essential test method components of the original and updated ECVAM Performance Standards (PS) (ECVAM, 2009a, b; ECVAM 2007). Moreover, the method's data relating to identification of non-classified vs classified substances formed the main basis for defining the specificity and sensitivity values of the original ECVAM PS.	<ul style="list-style-type: none"> • Fentem et al., 2001 • Portes et al., 2002 • Zuang et al., 2002 • Cotovio et al., 2005 • Hoffmann, 2006 • Spielmann et al., 2007 • Cotovio et al., 2007 • Eskes et al., 2007 • ECVAM, 2007a • ECVAM, 2007b • ECVAM, 2008 • ECVAM 2009 • ECVAM 2009a • ECVAM 2009b • Griesinger et al., 2010
2	EpiDerm™ SIT (EPI-200)	Initially the test method underwent full prospective validation together with Nr. 1. from 2005-2007. The test method components of this method were used to define the essential test methods components of the original and updated ECVAM PS (ECVAM, 2009a, b; ECVAM 2007).	<ul style="list-style-type: none"> • Fentem et al., 2001 • Zuang et al., 2002 • Kandarova et al., 2004 • Kandarova et al., 2005 • Hoffmann, 2006 • Spielmann et al., 2007 • Eskes et al., 2007 • ECVAM 2007a • ECVAM 2007b • ECVAM 2009 • ECVAM 2009a • ECVAM 2009b • Griesinger et al., 2010
		A modification of the EpiDerm™ SIT (EPI-200) was validated using the original ECVAM PS (ECVAM 2007a) in 2008	<ul style="list-style-type: none"> • ECVAM 2007a • ECVAM 2008 • ECVAM 2009 • Griesinger et al., 2010
2. Methods validated on the basis of Performance Standards			
Nr.	Test method name	Validation study type	References
3	SkinEthic™ RHE	Validation study based on the original ECVAM Performance Standards (ECVAM 2009a) in 2008.	<ul style="list-style-type: none"> • ECVAM 2007 • ECVAM 2008 • ECVAM 2009 • Griesinger et al., 2010
4	LabCyte EPI-MODEL24 SIT	Validation study (2011-2012) based on the Performance Standards (PS) of OECD TG 439 which are based on the updated ECVAM PS (ECVAM 2009a, b).	<ul style="list-style-type: none"> • ECVAM 2009a • ECVAM 2009b • Katoh et al., 2009 • Katoh et al., 2011 • OECD, 2011 • OECD, 2011a • Kojima et al, 2012

3.4.3 Non-standard test methods (not yet validated / accepted as OECD TGs)

Three test submissions of skin irritation test methods, all of which use simple reconstructed epidermis as the test system, were submitted to (EURL) ECVAM during 2009-2012 (see Appendix 2). Two of these methods have already been evaluated by ECVAM with respect to their similarity to validated reference methods. Based on a comparison of the essential test method components and the protocol parameters with the ECVAM Performance Standards (ECVAM, 2009d), it was established that these two methods are sufficiently similar to qualify for PS-based validation. The third method is still in the development phase.

One of the test methods is based on an "open source" concept. This test method was originally developed by Poumay et al. and has been published without any restrictions (Poumay et al., 2004). The "open source" concept would carry this further: all necessary information on both the production and maintenance of OS-Rep, i.e. the relevant Standard Operating Procedures (SOPs), will not be subject to any restrictions (e.g. patents or other intellectual property rights) as is the case with the current tissue models marketed by the various producers. The idea is that all interested users (either in academia, in research or contract-research organisations) will be able to reconstruct the RhE test system in their facilities and use the test system either in association with the relevant protocol for skin irritation testing (i.e. as skin irritation test method) or for other purposes (e.g. research and development). This approach is novel to an area of complex tissue-engineered test systems which, so far, are all manufactured and quality-controlled by the original test method producers who sell batch-controlled tissue kits with elements that are protected by intellectual property rights. The transferability of the reconstructed model, and the implementation of appropriate batch quality control procedures, need to be considered with care.

3.4.4 Quantitative Structure-Activity Relationships

Several QSAR modelling studies have attempted to model skin irritation or corrosion, focusing either on the prediction of tissue scores such as the primary irritation index (PII) or regulatory classifications. The status of literature-based models has changed little since the previous JRC reviews (Gallegos Saliner et al., 2006, 2008), with relatively few publications appearing since 2008 (e.g. Golla et al., 2009; Mombelli, 2008).

The simplest model for skin corrosion is based on pH measurements and takes the form of a simple decision rule:

If $\text{pH} < 2$ or $\text{pH} > 11.5$ then predict to be corrosive

This model is included in the OECD testing strategy for skin irritation and corrosion (OECD, 2002). Measurements of pH can also be used as a predictor of skin irritation (Worth & Cronin, 2001b). For defined chemical classes, QSARs have been reported for discriminating between corrosives and non-corrosives (Barratt 1996a, 1996b), and between irritants and non-irritants (Smith et al, 2000a, 2000b). A few models for predicting PII have also been reported, mainly for specific chemical classes (Hayashi et al., 1999; Kodithala et al., 2002), although Golla et al. (2009) reported a QSAR for predicting the PII of diverse chemical classes.

Software tools with functionalities for predicting skin irritation potential include Toxtree, the OECD QSAR Toolbox, Derek Nexus, TOPKAT, HazardExpert, Molcode QSARModel, and Multi-CASE. A comparison of the predictive capacities of three commercial tools, Derek, HazardExpert and TOPKAT (Mombelli, 2008) concluded that TOPKAT had the best performance (in terms of highest sensitivity). The freely available Toxtree and the OECD Toolbox tools include the so-called BfR rulebase, developed by the German Federal Institute for Risk Assessment (BfR; Bundesinstitut für Risikobewertung), which applies a set of physicochemical rules to predict the absence of effects, along with a set of structural rules to predict the presence of effects.

Evaluations of the predictive performance of the physicochemical (Rorije & Hulzebos, 2005) and structural (Saliner et al., 2007) rules found a high negative predictivity (99% for non-corrosives; 97% for non-irritancy) of the physicochemical exclusion rules, and a high positive predictivity of the structural inclusion rules (95% for corrosives; and 68% for irritants). These evaluations were performed before the BfR rulebase was implemented in Toxtree and the OECD Toolbox, so it would be worthwhile to revisit the analysis, using the more modern tools and recently published data. Both rules would need to be fulfilled in order to predict that a substance is non-irritant, i.e. a substance should fulfil the exclusion rules, and not have any alert for the inclusion rules. A limitation of this system is that the BfR exclusion rules depend on physicochemical properties of the substance, such as surface tension or lipid solubility, for which there is no reliable calculation method either in the Toolbox or in Toxtree. Hence, to use this system with some confidence, it would be necessary to have the measured values of the relevant physicochemical properties, unlike for the above-mentioned commercial tools which can be applied using only a structure as an input. The alerts from the inclusion rules might potentially be used on their own for a positive prediction, since this would represent a conservative approach.

The above-mentioned software tools can be used in a WoE approach or tiered testing strategy. As with all computational tools, the predictions needed to be carefully evaluated using information on the model characteristics. For classification and labelling, the BfR rulebase provides information that is closest to the regulatory goal, since the system was designed to predict former EU Risk Phrases for irritation (R38) and corrosion (R34, R35).

3.5 Status of Integrated Testing Strategies

There are two major integrated strategies available for the testing and evaluation of skin corrosion and skin irritation. Both make use of a wide range of information including from *in vitro* and *in vivo* test methods. In addition, a tiered approach for classification and labelling is described in UN GHS (UN, 2013). The EU CLP Regulation makes reference to this tiered approach (EU, 2008).

3.5.1 Testing and Evaluation Strategy in OECD TG 404

TG404 on the traditional Draize animal test contains a supplement which provides a "Testing and evaluation strategy for dermal irritation / corrosion". Importantly, unlike the body text of the TG, the supplement does not fall under the provisions of Mutual Acceptance of Data. This supplement was annexed to the TG in 2002 in relation to animal welfare concerns. This supplement outlines a possible "Testing and Evaluation

Strategy for Dermal Irritation/Corrosion", developed at an OECD workshop in 1996 (OECD, 1996). This strategy has been evaluated in relation to its ability to classify skin corrosives (Worth et al., 2008) and irritants (Hoffmann et al., 2008). The strategy foresees a sequential approach based on the use of:

- (a) available information as well structural and physicochemical data;
- (b) validated and accepted *in vitro* / *ex vivo* tests for skin corrosion and irritation;
- (c) animal testing (i.e. traditional Draize rabbit test).

The strategy contains eight steps that can be grouped according to the three above-mentioned data sources:

(a) *Available information (human or animal data & structural and physicochemical information)*

1. Consideration of existing human or animal data with regard to skin corrosion / irritation.
2. Structure-Activity Relationship (SAR) data with regard to skin corrosion / irritation.
3. pH measurements: extreme pH (≤ 2 or ≥ 11.5) suggests corrosive properties
4. systemic toxicity data via the dermal route (can also be considered before steps 2 and 3)

(b) *In vitro tests*

5. conducting validated and accepted *in vitro* or *ex vivo* test for skin corrosion (consider positives)
6. conducting validated and accepted *in vitro* or *ex vivo* test for skin irritation (consider positives)

(c) *In vivo rabbit test*

7. conducting TG404 test using one animal to test for corrosive effects
8. if not corrosive in step 7: perform testing using one or two animals to assess whether substance is corrosive or irritant or not.

This strategy is not currently (October 2013) in line with the status of acceptance of alternative methods as it foresees only the use of positive results from *in vitro* methods for skin corrosion and irritation and suggests confirmatory testing for negative *in vitro* results in the animal test. However, *in vitro* methods, especially those based on Reconstructed human Epiderms (RhE) applicable to a wide range of chemicals, can be used to conclude both on the presence and absence of skin corrosion potential as well as on the presence and absence of skin irritation potential – at least in areas (such as the EU) not implementing optional Category 3 ("mild irritants") and not corrosive subcategories (1A, 1B, 1C). Currently, in the context of the OECD expert group on skin irritation/corrosion, there are efforts to update this testing strategy and provide sufficient guidance for its proper use. This activity is led by Germany. The resulting guidance document uses the Integrated Testing Strategy (ITS) developed for REACH (ECHA, 2012). As a new element, a more consistent and explicit description of the performance characteristics of the individual elements of the ITS has been included, providing, for example, for each test method information on the predictive capacity, reliability, applicability and limitations. Initially the derivation of possible weighing factors attached to the various data sources was also considered but this was not followed up. This would have included the definition of a consistent data matrix of chemicals with credible reference data and an analysis of the relative contribution of each data source to the decision making using a semi-quantitative approach. The derived factors would have been used as weighing factors for the individual test methods when conducting a Weight of Evidence (WoE) analysis of chemicals.

3.5.2 Integrated Testing Strategy in REACH guidance

The testing strategy recommended in REACH guidance (ECHA, 2012) is based on expert knowledge and judgement, and consists of three general steps:

STEP (1) Gathering and evaluation of existing information, taking into account data on/from:

- pH
- physicochemical properties
- human data
- existing animal data from animal corrosion/irritation studies (including from chemical analogues)
- dermal toxicity data
- other studies (sensitisation and repeated exposure)
- (Q)SAR
- Read across or grouping
- Validated *In vitro* corrosion test methods
- Validated *In vitro* irritation test methods
- Non validated *in vitro* assays

STEP (2) Weight of Evidence integration of all information gathered in Step (1)

STEP (3) New testing if the information integrated/evaluated in Step (2) is found insufficient to make a decision on classification and labelling (goal one of the human health hazard assessment under REACH, REACH Annex I):

- skin corrosion *in vitro* data from an OECD adopted test method
- skin irritation *in vitro* data from an OECD adopted test method
- skin irritation *in vitro* data from a non-validated *in vitro* test
- *in vivo* skin irritation data (TG404)

3.5.3 United Nations tiered approach for classification and labelling

The Globally Harmonised System for Classification and Labelling includes a "decision logic" for the classification and labelling of skin corrosion / irritation (UN, 2013). Formally, the decision logic is based on existing information only. The UN GHS decision logic is currently undergoing revision in the context of the 5th edition of UN GHS. The "tiered evaluation for skin corrosion and irritation" provides guidance on "how to organize existing information on a substance and to make a weight of evidence decision about hazard assessment and hazard classification (ideally without conducting new animal tests)" (UN, 2013). This tiered approach is referenced in the EU CLP Regulation (EU, 2008).

3.6. Conclusions

Skin corrosion and irritation are important health endpoints due to the readiness with which skin can be exposed to chemicals and as evidenced by the high incidence of Irritant Contact Dermatitis (ICD) compared with other health endpoints. Hazard information on these endpoints are required under REACH and the Biocidal Products Regulation, and can also be used in the implementation of the CLP Regulation.

The mechanisms of skin corrosion are relatively simple and largely based on the chemical destruction / erosion of the skin tissue and subsequent irreversible necrosis of

the affected tissue. In contrast, the mechanisms of skin irritation are more complex. While the triggers of irritation can be assumed to be largely based on mechanisms inducing tissue trauma (e.g. reactive chemicals, chemicals interfering with the integrity of plasma membranes), the subsequent events involving the innate immune system and nerve endings of the skin leading to a local inflammatory reaction (the actual irritation response) are complex and not yet fully understood. Nevertheless, there is sufficient understanding of the key events underlying skin irritation to develop a putative AOP (Appendix 1). Furthermore, by modelling up-stream key events rather than mid-stream inflammatory events, *in vitro* methods have been developed that can predict skin irritation potential with great accuracy. QSAR models have been developed mainly for predicting the adverse outcome (skin corrosion or irritation) itself, rather than intermediate events. These approaches are applicable to specific chemical classes and can provide useful information in the context of integrated approaches. Further QSAR efforts should focus on the modelling of specific key events in the AOP by exploiting available datasets generated by standardised *in vitro* methods.

When limitations/scope of the accepted *in vitro* methods are considered, these methods can be used to meet the information requirement of REACH Regulation, provided that the testing strategy specified above is followed. Guidance on this will be available from ECHA (<http://echa.europa.eu/support/testing-methods-and-alternatives>). Technically speaking, information from *in vitro* test methods needs to be submitted as part of a Weight of Evidence approach under REACH; the respective instructions are given in http://echa.europa.eu/documents/10162/13655/pg_report_in_vitro_data_en.pdf

For both skin corrosion and irritation, a number of different testing and non-testing methods are available in addition to the standard animal test. *In vitro* methods have undergone full prospective validation or validation based on performance standards and are internationally recognized (OECD Test Guidelines) and accepted by EU regulatory authorities (EU test methods regulation) as standard test methods. Detailed Performance Standards are available for all *in vitro* methods (corrosion and irritation) and in particular for skin models based on Reconstructed human Epidermis (RhE). Moreover, testing strategies for skin corrosion and irritation have been published by regulatory bodies. These provide guidance on how to combine information on physicochemical properties and QSARs with information from *in vitro* methods, the *in vivo* test and, where available, existing human information to support decisions on classification and labelling.

3.7 References

- Al-Jaberi H, Marks R (1984). Studies of the clinically uninvolved skin in patients with dermatitis. *Br J Dermatol*. 111(4): 437-43.
- Astner S, González S, Gonzalez E (2006). Noninvasive evaluation of allergic and irritant contact dermatitis by in vivo reflectance confocal microscopy. *Dermatitis* 17(4): 182-91.
- Atherton DJ (2001). The aetiology and management of irritant diaper dermatitis. *J Eur Acad Dermatol Venereol*. 15 Suppl 1:1-4.
- Barratt MD (1996a). Quantitative structure-activity relationships (QSARs) for skin corrosivity of organic acids, bases and phenols: Principal components and neural network analysis of extended datasets. *Toxicology in Vitro* 10(1): 85-94.
- Barratt MD (1996b). Quantitative structure-activity relationships for skin irritation and corrosivity of neutral and electrophilic organic chemicals. *Toxicology in Vitro* 10: 247-256.
- Barratt MD, Brantom PG, Fentem JH, Gerner I, Walker AP, Worth AP (1998). The ECVAM international validation study for in vitro tests for skin corrosivity. 1. Selection and distribution of the test chemicals. *Toxicology in Vitro* 12: 471-482.
- Bartek MJ, LaBudde JA, Maibach HI (1972). Skin permeability in vivo: comparison in rat, rabbit, pig and man. *J Invest Dermatol*. 58(3): 114-23.
- Basketter DA, Chamberlain M, Griffiths HA, Rowson M, Whittle E, York M (1997). The classification of skin irritants by human patch test. *Food Chem Toxicol*. 35(8):845-52.
- Basketter DA, Marriott M, Gilmour NJ, White IR (2004a). Strong irritants masquerading as skin allergens: the case of benzalkonium chloride. *Contact Dermatitis* 50(4): 213-7.
- Basketter DA, York M, McFadden JP, Robinson MK (2004b). Determination of skin irritation potential in the human 4-h patch test. *Contact Dermatitis* 51(1):1-4.
- Basketter DA, Jirova D, Kandarova H (2012). Review of skin irritation/corrosion hazards on the basis of human data: a regulatory perspective. *Interdiscip Toxicol* 5(2): 98-104
- Beresford L, Orange O, Bell EB, Miyan JA (2004) Nerve fibres are required to evoke a contact sensitivity response in mice. *Immunology* 111: 118-125.
- Botham PA, Chamberlain M, Barratt MD, Curren RD, Esdaile DJ, Gardner JR, Gordon VC, Hildebrand B, Lewis RW, Liebsch M, Logemann P, Osborne R, Ponc M, Regnier JF, Steiling W, Walker AP, Balls M (1995). A Prevalidation study on in vitro skin corrosivity testing. The report and recommendations of ECVAM workshop 6. *ATLA* 23: 219-255.
- Cork MJ, Robinson DA, Vasilopoulos Y, Ferguson A, Moustafa M, MacGowan A, Duff GW, Ward SJ, Tazi-Ahnini R (2006). New perspectives on epidermal barrier dysfunction in atopic dermatitis: gene-environment inter-actions. *J Allergy Clin Immunol*. 118(1): 3-21.
- Cotovio J, Grandidier MH, Lelièvre D, Roguet R, Tinois-Tessonnaud E and Leclaire J (2007). In vitro acute skin irritancy of chemicals using the validated EPISKIN model in a tiered strategy - Results and performances with 184 cosmetic ingredients. *AATEX* 14, 351-358.
- Chew AL, Maibach HI (2006). *Irritant Dermatitis*. Berlin, Springer.
- Cotovio J, Grandidier MH, Portes P, Roguet R and Rubinsteen G (2005). The in vitro acute skin irritation of chemicals: optimisation of the EPISKIN prediction model within the framework of the ECVAM validation process. *ATLA* 33: 329-349.
- Cowley NC, Farr PM (1992). A dose-response study of irritant reactions to sodium lauryl sulphate in patients with seborrhoeic dermatitis and atopic eczema. *Acta Derm Venereol*. 72(6):432-5.

Dereure O (2007). Filaggrin and genodermatosis: role in ichthyosis vulgaris and in atopic dermatitis. *Ann Dermatol Venereol.*134(4 Pt 1): 416.

Dickel H, Kuss O, Schmidt A, Kretz J, Diepgen TL (2002). Importance of irritant contact dermatitis in occupational skin disease. *Am J Clin Dermatol.* 3(4): 283-9.

Dahl MV (1988). Chronic, irritant contact dermatitis: mechanisms, variables, and differentiation from other forms of contact dermatitis. *Adv Dermatol.*3: 261-75.

Draize JH, Woodard G & Calvery HO (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *J. Pharmacol. and Exp. Therapeutics* 82, 377–390.

Eberhard Y, Ortiz S, Ruiz Lascano A, Kuznitzky R, Serra HM (2004). Up-regulation of the chemokine CCL21 in the skin of subjects exposed to irritants. *BMC Immunol.* 26: 5-7.

ECVAM (1998). Statement on the scientific validity of the EpiSkin test (an in vitro test for skin corrosivity), issued by the ECVAM Scientific Advisory Committee (ESAC10), 3 April 1998. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2000a). Statement on the application of the EpiDerm human skin model for skin corrosivity testing, issue by the ECVAM Scientific Advisory Committee (ESAC14), 21 March 2000. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2000b). Statement on the application of the Corrositex assay for skin corrosivity testing, issue by the ECVAM Scientific Advisory Committee (ESAC15), 5-6 December 2000. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2006). Statement on the application of the SkinEthic™ human skin model for skin corrosivity testing, issued by the ECVAM Scientific Advisory Committee (ESAC25), 17 November 2006. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2007a). Statement on the validity of in vitro tests for skin irritation, issued by the ECVAM Scientific Advisory Committee (ESAC26), 27 April 2007. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2007b). Performance Standards for applying human skin models to in vitro skin irritation testing. Available from: http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing

N.B. These are the original PS used for the validation of two test methods. These PS should not be used any longer as an updated version (8) is now available.

ECVAM (2008). Statement on the scientific validity of in vitro tests for skin irritation testing, issued by the ECVAM Scientific Advisory Committee (ESAC29), 5 November 2008. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2009a). Performance Standards for in vitro skin irritation test methods based on Reconstructed human Epidermis (RhE) Available from: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/validation-regulatory-acceptance/topical-toxicity/skin-irritation

ECVAM (2009b), ESAC Statement on the Performance Standards (PS) for in vitro skin irritation testing using Reconstructed human Epidermis, issued by the ECVAM Scientific Advisory Committee (ESAC31), 8 July 2009. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2009c). ESAC statement on the scientific validity of an in-vitro test method for skin corrosivity testing: the EST-1000, issued by the ECVAM Scientific Advisory Committee (ESAC30), 12 June 2009. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECVAM (2009d). Statement on the “Performance under UN GHS of three in vitro assays for skin irritation testing and the adaptation of the Reference Chemicals and Defined Accuracy Values of the ECVAM skin irritation Performance Standards”, issued by the ECVAM Scientific Advisory Committee (ESAC30), 9 April 2009. Online: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions

ECETOC (1995). European Centre for Ecotoxicology and Toxicology of Chemicals. TR 066 - Skin Irritation and Corrosion: Reference Chemicals Data Bank March 1995. Online: <http://www.ecetoc.org/technical-reports>.

ECHA (2012). Guidance on information requirements and chemical safety assessment. Chapter R.7a: Endpoint specific guidance. In "Guidance for the implementation of REACH". Version 2.0. November 2012. Online: http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf

Elsner P (1994). Irritant dermatitis in the workplace. *Dermatol Clin.*12(3):461-7.

Eskes C, Detappe V, Koëter H, Kreysa J, Liebsch M, Zuang V, Amcoff P, Barroso J, Cotovio J, Guest R, Hermann M, Hoffmann S, Masson P, Alépée N, Arce LA, Brüscheweiler B, Catone T, Cihak R, Clouzeau J, D'Abrosca F, Delveaux C, Derouette JP, Engelking O, Facchini D, Fröhlicher M, Hofmann M, Hopf N, Molinari J, Oberli A, Ott M, Peter R, Sá-Rocha VM, Schenk D, Tomicic C, Vanparys P, Verdon B, Wallenhorst T, Winkler GC, Depallens O (2012). Regulatory assessment of in vitro skin corrosion and irritation data within the European framework: Workshop recommendations. *Regul Toxicol Pharmacol.* 62(2):393-403.

Eskes C, Cole T, Hoffmann S, Worth A, Cockshott A, Gerner I and Zuang V (2007). The ECVAM international validation study on in vitro tests for acute skin irritation: selection of test chemicals. *ATLA* 35, 603-619.

EC (2006). Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC

EC (2008a). Council Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH)

EC (2008b). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006.

EC (2009). Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products

EU (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products

EU (2013). Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.

- Fartasch M (1997). Ultrastructure of the epidermal barrier after irritation. *Microsc Res Tech.* 37(3):193-9.
- Fentem JH, Archer GE, Balls M, Botham PA, Curren RD, Earl LK, Esdaile DJ, Holzhütter HG, Liebsch M (1998). The ECVAM International Validation Study on In Vitro Tests for Skin Corrosivity. 2. Results and Evaluation by the Management Team. *Toxicol In Vitro.* 12(4):483-524.
- Fentem JH, Briggs D, Chesné C, Elliot GR, Harbell JW, Heylings JR, Portes P, Roguet R, van de Sandt JJM and Botham P (2001). A prevalidation study on in vitro tests for acute skin irritation, Results and evaluation by the Management Team. *Toxicol. in Vitro* 15, 57-93.
- Fluhr JW, Kuss O, Diepgen T, Lazzerini S, Pelosi A, Gloor M, Berardesca E (2001). Testing for irritation with a multifactorial approach: comparison of eight non-invasive measuring techniques on five different irritation types. *Br J Dermatol.* 145(5): 696-703.
- Fluhr JW, Bankova L, Fuchs S, Kelterer D, Schliemann-Willers S, Norgauer J, Kleesz P, Grieshaber R, Elsner P (2004). Fruit acids and sodium hydroxide in the food industry and their combined effect with sodium lauryl sulphate: controlled in vivo tandem irritation study. *Br J Dermatol.* 151(5): 1039-48.
- Fluhr JW, Kelterer D, Fuchs S, Kaatz M, Grieshaber R, Kleesz P, Elsner P (2005a). Additive impairment of the barrier function and irritation by biogenic amines and sodium lauryl sulphate: a controlled in vivo tandem irritation study. *Skin Pharmacol Physiol.* 18(2): 88-97.
- Fluhr JW, Akengin A, Bornkessel A, Fuchs S, Praessler J, Norgauer J, Grieshaber R, Kleesz P, Elsner P. (2005b) Additive impairment of the barrier function by mechanical irritation, occlusion and sodium lauryl sulphate in vivo. *Br J Dermatol.* 153(1): 125-31.
- Fluhr JW, Praessler J, Akengin A, Fuchs SM, Kleesz P, Grieshaber R, Elsner P. (2005c) Air flow at different temperatures increases sodium lauryl sulphate-induced barrier disruption and irritation in vivo. *Br J Dermatol.* 152(6): 1228-34.
- Fluhr JW, Darlenski R, Angelova-Fischer I, Tsankov N, Basketter D (2008). Skin irritation and sensitization: mechanisms and new approaches for risk assessment. 1. Skin irritation. *Skin Pharmacol Physiol.* 21(3):124-35.
- Gallegos Saliner A, Patlewicz G & Worth AP (2006). Review of Literature-Based Models for Skin and Eye Irritation and Corrosion. JRC Report EUR 22320 EN. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/QSAR_Review_Irritation.pdf
- Gallegos Saliner A, Patlewicz G & Worth AP (2008). A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. *QSAR & Combinatorial Science* 27: 49-59.
- Golla S, Madihally S, Robinson RL & Gasem KAM (2009). Quantitative structure-property relationships modeling of skin irritation. *Toxicology in Vitro:* 23: 176–84.
- Griesinger C, Barroso J, Zuang V, Cole T, Genschow E and Liebsch M (2010) Explanatory background document to the OECD Test Guideline on in vitro skin irritation testing. OECD Series on Testing and Assessment, No. 137, OECD, Paris. An updated version is available from http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/validation-regulatory-acceptance/topical-toxicity/skin-irritation
- Griesinger C, Zuang V (2012). Skin Irritation Corrosion. In: *ECVAM and New Technologies for Toxicity Testing in New Technologies for Toxicity Testing*, eds Balls M, Combes RD, Bhogal N. *Adv Exp Med Biol.*
- Hachem JP, Crumrine D, Fluhr J, Brown BE, Feingold KR, Elias PM (2003). pH directly regulates epidermal permeability barrier homeostasis, and stratum corneum integrity/cohesion. *J Invest Dermatol.* 121(2): 345-53.
- Hayashi M et al. (1999). A quantitative structure-activity relationship study of the skin irritation potential of phenols. *Toxicology in Vitro* 13: 915–922.

- Hoefakker S, Caubo M, van 't Erve EH, Roggeveen MJ, Boersma WJ, van Joost T, Notten WR, Claassen E (1995). In vivo cytokine profiles in allergic and irritant contact dermatitis. *Contact Dermatitis* 33(4): 258-66.
- Hoffmann J, Heisler E, Karpinski S, Losse J, Thomas D, Siefken W, Ahr HJ, Vohr HW, Fuchs HW (2005). Epidermal-skin-test 1,000 (EST-1,000)-a new reconstructed epidermis for in vitro skin corrosivity testing. *Toxicol In Vitro* 19(7): 925-9.
- Hoffmann S (2006). ECVAM skin irritation validation study phase II: Analysis of the primary endpoint MTT and the secondary endpoint IL1- α . Available from: http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing
- Hoffmann S, Saliner AG, Patlewicz G, Eskes C, Zuang V & Worth AP (2008). A feasibility study developing an integrated testing strategy assessing skin irritation potential of chemicals. *Toxicology Letters* 180: 9–20.
- Houben E, De Paepe K, Rogiers V (2007). A keratinocyte's course of life. *Skin Pharmacol Physiol.* 20(3): 122-32.
- ICCVAM (1999). Interagency Coordinating Committee on the Validation of Alternative Methods. Corrositex. An In Vitro Test Method for Assessing Dermal Corrosivity of Chemicals. NIH Publication No. 99-4495. http://iccvam.niehs.nih.gov/docs/dermal_docs/corprprep.pdf
- Jírová D, Basketter D, Liebsch M, Bendová H, Kejlová K, Marriott M, Kandárová H (2010). Comparison of human skin irritation patch test data with in vitro skin irritation assays and animal data. *Contact Dermatitis* 62(2): 109-16.
- Kandárová H, Liebsch M, Spielmann H, Genschow E, Schmidt E, Traue D, Guest R, Whittingham A, Warren N, Gamer AO, Remmele M, Kaufmann T, Wittmer E, De Wever B, Rosdy M (2006). Assessment of the human epidermis model SkinEthic RHE for in vitro skin corrosion testing of chemicals according to new OECD TG 431. *Toxicology In Vitro* 20(5): 547-59.
- Kandárová H, Liebsch M, Genschow E, Gerner I, Traue D, Slawik B and Spielmann H (2004). Optimisation of the EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests. *ALTEX* 21: 107–114.
- Kandárová H, Liebsch M, Gerner I, Schmidt E, Genschow E, Traue D and Spielmann H (2005). The EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests – An assessment of the performance of the optimised test. *ATLA* 33, 351-367.
- Kao J, Patterson FK, Hall J (1985). Skin penetration and metabolism of topically applied chemicals in six mammalian species, including man: an in vitro study with benzo[a]pyrene and testosterone. *Toxicol Appl Pharmacol.* 81(3 Pt 1): 502-16.
- Katoh M and Hata K (2011). Refinement of LabCyte EPI-MODEL24 skin Irritation test method for adaptation to the requirements of OECD test guideline 439. *AATEX*, 16, 111-122.
- Katoh M, Hamajima F, Ogasawara T and Hata K (2009). Assessment of human epidermal model LabCyte EPI-MODEL for in vitro skin irritation testing according to European centre for the validation of alternative methods (ECVAM)-Validated Protocol. *J Toxicol Sci* 34, 327-334.
- Kindt TJ, Osborne BA, Goldsby RA (2006). *Immunology (The "Kuby")*, 6th edition. Freeman and Company, New York.
- Kodithala K, Hopfinger AJ, Thompson ED & Robinson MK (2002). Prediction of skin irritation from organic chemicals using membrane-interaction QSAR analysis. *Toxicological Sciences* 66, 336–346.
- Kojima H, Ando Y, Idehara K, Katoh M, Kosaka T, Miyaoka E, Shinoda S, Suzuki T, Yamaguchi Y, Yoshimura I, Yuasa A, Watanabe Y and Omori T (2012). Validation Study of the In Vitro Skin Irritation Test with the LabCyte EPI-MODEL24. *ATLA* 40, 33-50.

Kolle SN, Sullivan KM, Mehling A, van Ravenzwaay B, Landsiedel R (2013). Erratum to "Applicability of in vitro tests for skin irritation and corrosion to regulatory classification schemes: substantiating test strategies with data from routine studies" [Regul. Toxicol. Pharmacol. (2012) 402-414]. Regul Toxicol Pharmacol. 65(3):366-78.

Lammintausta K, Maibach HI, Wilson D (1988). Mechanisms of subjective (sensory) irritation. Propensity to non-immunologic contact urticaria and objective irritation in stingers. Derm Beruf Umwelt. 36(2):45-9.

Liebsch M, Traue D, Barrabs C, Spielmann H, Uphill P, Wilkins S, McPherson JP, Wiemann C, Kaufmann T, Remmele M, Holzhütter HG (2000). The ECVAM prevalidation study on the use of EpiDerm for skin corrosivity testing. ATLA 28, pp. 371-401.

Mizutani H, Schechter N, Lazarus G, Black RA, Kupper TS (1991). Rapid and specific conversion of precursor interleukin 1 beta (IL-1 beta) to an active IL-1 species by human mast cell chymase. J Exp Med. 174(4):821-5.

Mombelli E (2008). An evaluation of the predictive ability of the QSAR software packages, DEREK, HAZARDEXPERT and TOPKAT, to describe chemically-induced skin irritation. ATLA 36:15-24.

Nicolaidis N (1964). Human skin surface lipids-origins, composition and possible function. In: Advances in Biology of Skin. 4: The Sebaceous Glands 1963 (ed. by Montagna, W., Ellis, R. A., and Silver, A. F.). Pergamon Press, Oxford, pp 167-187.

OECD (1996). Final Report on the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Held on Solna, Sweden, 22-24 January 1996

Note that the link provided in TG404 <http://www1.oecd.org/ehs/test/background.htm> does not work (attempted access July 2013) and the document seems not retrievable at present.

OECD (2002). Test No. 404: Acute Dermal Irritation/Corrosion, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing.

OECD (2011).. Validation report for the skin irritation test method using LabCyte EPI-MODEL24, OECD Series on Testing and Assessment No. 159, OECD, Paris. Available from: <http://www.oecd.org/env/testguidelines>

OECD (2011a). Peer review report of validation of the skin irritation test using LabCyte EPI-MODEL24, OECD Series on Testing and Assessment No. 155, OECD, Paris. Available from: <http://www.oecd.org/env/testguidelines>

OECD (2012). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168. Available online at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2012\)10/PART1&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En)

Phillips L 2nd, Steinberg M, Maibach HI, Akers WA (1972). A comparison of rabbit and human skin response to certain irritants. Toxicol Appl Pharmacol. 21(3):369-82.

Ponec M, Kempenaar J (1995). Use of human skin recombinants as an in vitro model for testing the irritation potential of cutaneous irritants. Skin Pharmacol. 8(1-2):49-59.

Portes P, Grandidier MH, Cohen C, Roguet R (2002). Refinement of the EpiSkin protocol for the assessment of acute skin irritation of chemicals: follow-up to the ECVAM prevalidation study. Toxicol In Vitro. 16(6):765-70.

Poumay Y, Dupont F, Marcoux S, Leclercq-Smekens M, Hérin M, Coquette A (2004). A simple reconstructed human epidermis: preparation of the culture model and utilization in in vitro studies. Arch Dermatol Res. 296(5):203-11.

Prudovsky I, Mandinova A, Soldi R, Bagala C, Graziani I, Landriscina M, Tarantini F, Duarte M, Bellum S, Doherty H, Maciag T (2003). The non-classical export routes: FGF1 and IL-1alpha point the way. *J Cell Sci.* 116(Pt 24): 4871-81.

Prudovsky I, Tarantini F, Landriscina M, Neivandt D, Soldi R, Kirov A, Small D, Kathir KM, Rajalingam D, Kumar TK (2008). Secretion without Golgi. *J Cell Biochem.* 103(5):1327-43.

Rorije, E. & Hulzebos, E. (2005). Evaluation of (Q)SARs for the prediction of Skin Irritation/Corrosion Potential. Physicochemical exclusion rules. Final report for ECB contract IHCP.B430206. Joint Research Centre, Ispra, Italy. Available from:

http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/Evaluation_of_Eye_Irritation_Q_SARs.pdf

Saeki H, Moore AM, Brown MJ, Hwang ST (1999). Cutting edge: secondary lymphoid-tissue chemokine (SLC) and CC chemokine receptor 7 (CCR7) participate in the emigration pathway of mature dendritic cells from the skin to regional lymph nodes. *J Immunol.* 162(5):2472-5.

Saliner AG, Tsakovska I, Pavan M, Patlewicz G & Worth AP (2007). Evaluation of SARs for the prediction of skin irritation/corrosion potential: structural inclusion rules in the BfR decision support system. *SAR and QSAR in Environmental Research* 18:331–342.

Sato K, Sugibayashi K, Morimoto Y (1991). Species differences in percutaneous absorption of nicorandil. *J Pharm Sci.* 80(2):104-7.

Schaefer H, Filaquier C (1992). Skin metabolism. *Pathol Biol (Paris)*. Feb;40(2):196-204. Review (French).

Schmid-Wendtner MH, Korting HC (2006). The pH of the skin surface and its impact on the barrier function. *Skin Pharmacol Physiol.* 19(6):296-302.

Smith JS, Macina OT, Sussman NB, Luster MI, Karol MH (2000a). A robust structure-activity relationship (SAR) model for esters that cause skin irritation in humans. *Toxicological Sciences* 55, 215-222.

Smith JS, Macina OT, Sussman NB, Karol MH, Maibach HI (2000b). Experimental validation of a structure-activity relationship model of skin irritation by esters. *Quantitative Structure-Activity Relationships* 19, 467–474.

Snater E, Janssen EA, van der Valk PG, van de Kerkhof PC (1995). Transepidermal water vapour loss is not increased during and following dithranol irritation. *Br J Dermatol.* 132(6):908-12.

Spiekstra SW, Toebak MJ, Sampat-Sardjoepersad S, van Beek PJ, Boorsma DM, Stoof TJ, von Blomberg BM, Scheper RJ, Bruynzeel DP, Rustemeyer T, Gibbs S (2005). Induction of cytokine (interleukin-1alpha and tumor necrosis factor-alpha) and chemokine (CCL20, CCL27, and CXCL8) alarm signals after allergen and irritant exposure. *Exp Dermatol.* 14(2):109-16.

Spielmann H, Hoffmann S, Liebsch M, Botham P, Fentem JH, Eskes C, Roguet R, Cotovio J, Cole T, Worth A, Heylings J, Jones P, Robles C, Kandárová H, Gamer A, Remmele M, Curren R, Raabe H, Cockshott A, Gerner I, Zuang V (2007). The ECVAM international validation study on in vitro tests for acute skin irritation: report on the validity of the EPISKIN and EpiDerm assays and on the Skin Integrity Function Test. *ATLA* 35(6):559-601.

Stamatas GN, Morello AP, Mays DA (2013). Early inflammatory processes in the skin. *Curr Mol Med.* 13(8):1250-69.

Steinsträsser I, Merkle HP (1995) Dermal metabolism of topically applied drugs: pathways and models reconsidered. *Pharm Acta Helv.* 70(1):3-24.

Sugibayashi K, Watanabe T, Hasegawa T, Takahashi H, Ishibashi T (2002). Kinetic analysis on the in vitro cytotoxicity using Living Skin Equivalent for ranking the toxic potential of dermal irritants. *Toxicology in Vitro* 16(6):759-63.

Tharp MD (1991) The mast cell and its mediators. In *Physiology, Biochemistry, and Molecular Biology of the Skin*. 2nd edition, edited by LA Goldsmith, pp. 1019-1083. Oxford University Press, New York.

Tornier C, Roquet M, Fraissinette AB (2010). Adaptation of the validated SkinEthic™ Reconstructed Human Epidermis (RHE) skin corrosion test method to 0.5 cm² tissue sample. *Toxicology In Vitro* 24: 1379-1385.

Turner S, Carder M, van Tongeren M, McNamee R, Lines S, Hussey L, Bolton A, Beck MH, Wilkinson M, Agius R (2007). The incidence of occupational skin disease as reported to The Health and Occupation Reporting (THOR) network between 2002 and 2005. *Br J Dermatol*. 157(4):713-22.

UN (2011), *Globally Harmonized System of Classification and Labelling of Chemicals*, 4th revised edition. United Nations, New York and Geneva.

UN (2013). *Globally Harmonized System of Classification and Labelling of Chemicals (GHS)*, 5th revised edition. United Nations, New York and Geneva.

http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf.

UN (2013). Report of the Committee of Experts on the Transport of Dangerous Goods and on the Globally Harmonized System of Classification and Labelling of Chemicals on its sixth session: Amendments to the fourth revised edition of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) (ST/SG/AC.10/30/Rev.4). Available from: <http://www.unece.org/fileadmin/DAM/trans/doc/2013/dgac10/ST-SG-AC10-40a3e.pdf>

Wasserman SI (1983). The mast cell and its mediators. In: *Biochemistry and Physiology of the skin*. Edited by LA Goldsmith, pp. 878-898. Oxford University Press, New York.

Wasserman SI (1990). Mast cell biology. *J Allergy Clin Immunol*. 86(4 Pt 2):590-3.

Weil CS, Scala RA (1971). Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. *Toxicol Appl Pharmacol*. 19(2): 276-360.

Wells T, Basketter DA, Schröder KR (2004). In vitro skin irritation: facts and future. State of the art review of mechanisms and models. *Toxicology in Vitro* 18, 231–243.

Welzel J, Bruhns M, Wolff HH (2003). Optical coherence tomography in contact dermatitis and psoriasis. *Arch Dermatol Res*. 295(2):50-5.

Wigger-Alberti W, Elsner P (2000). Contact Dermatitis due to Irritation, p99-110. In *Handbook of Occupational Dermatology*. Berlin, Springer.

Wigger-Alberti W, Spoo J, Schliemann-Willers S, Klotz A, Elsner P (2002). The tandem repeated irritation test: a new method to assess prevention of irritant combination damage to the skin. *Acta Derm Venereol*. 82(2):94-7.

Willis CM, Young E, Brandon DR, Wilkinson JD (1986). Immunopathological and ultrastructural findings in human allergic and irritant contact dermatitis. *Br J Dermatol*. 115(3):305-16.

Wood LC, Elias PM, Calhoun C, Tsai JC, Grunfeld C, Feingold KR (1996). Barrier disruption stimulates interleukin-1 alpha expression and release from a pre-formed pool in murine epidermis. *J Invest Dermatol*. 106(3):397-403.

Worth AP, Fentem JH, Balls M, Botham PA, Curren RD, Earl LK, Esdaile DJ & Liebsch M (1998). An evaluation of the proposed OECD testing strategy for skin corrosion. *ATLA* 26: 709-720.

Worth AP & Cronin MT (2001a). The use of bootstrap resampling to assess the variability of Draize tissue scores. *ATLA* 29: 557–573.

Worth AP & Cronin MTD (2001b). The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity. *Toxicology* 169,119-131.

Zhang Q, Grice JE, Wang G, Roberts MS (2009). Cutaneous metabolism in transdermal drug delivery. *Curr Drug Metab.* 10(3):227-35.

Zuang V, Balls M, Botham PA, Coquette A, Corsini E, Curren RD, Elliot GR, Fentem JH, Heylings JR, Liebsch M, Medina J, Roguet R, van De Sandt JJM, Wiemann C and Worth A (2002). Follow-up to the ECVAM prevalidation study on in vitro tests for acute skin irritation, The European Centre for the Validation of Alternative Methods Skin Irritation Task Force report 2. *ATLA* 30, 109-129.

Appendix to Chapter 3. Putative Adverse Outcome Pathway for Skin Irritation

General considerations:

- Skin irritation – in its clinical manifestation – is a result of inflammatory processes in the skin.
- Inflammation is a physiological response of any tissue to injury and trauma.
- Chemically-induced skin inflammation, as most inflammatory processes, is triggered by tissue trauma: impairment and damage of the various cell populations building up the affected tissue. It may however also involve more specific inflammatory activation of cells as well as neurogenic inflammatory pathways (i.e. activation of nerve endings in the dermal layers of the skin). This tissue trauma leads to an inflammatory reaction causing, as downstream effects, the classical inflammatory symptoms such as redness, swelling etc. Moreover, in case of "sensory"/"subjective" irritation, symptoms may be restricted to sensations of itching, stinging, burning and tingling without classical manifestations of irritation.
- Due to the complex interaction of the various key events (e.g. dermal penetration, destruction of skin barrier, damage of skin cells etc.) there is the possibility that chemicals show additive effects when applied in combination (e.g. in mixtures). Additive effects in skin irritation have been shown for instance for SLS, biogenic amines, solvents and fruit acids (Fluhr et al, 2004; Fluhr et al., 2005a; Wigger-Alberti et al, 2002).

Key events:

Key event 1 - Dermal bioavailability, skin penetration and absorption

It is self-evident that only chemicals that penetrate into the skin (epidermis and dermis) or are absorbed by the skin can trigger an inflammatory reaction in the skin tissue. The intrinsic physicochemical properties of chemicals (size, net charge, lipophilicity, partition coefficient (logP) etc.) are important determinants of their bioavailability. Moreover, the intrinsic skin condition is another important determinant for percutaneous absorption. This includes the integrity of the so-called "acid mantle" of the skin (the pH of skin ranges between 4.2 and 5.6) The acid mantle is based on a molecular film composed of triglycerides, phospholipids, esterified cholesterol released by holocrine sebaceous glands in the skin (Nikolaides, 1963). The skin pH appears to be a regulator of barrier function (Schmidt-Wendtner & Korting, 2006) by affecting epidermal permeability barrier homeostasis and stratum corneum cohesion/integrity (Hachem et al., 2003). The incomplete acid mantle in neonate skin is a key factor involved in diaper dermatitis (Atherton, 2001). Also in adults the acid mantle may vary due to endogenous (age, ethnicity, anatomical site, gender, other disorders) as well as exogenous factors (use of cosmetics, contact with specific chemicals (e.g. for occupational reasons), seasonal change). Other factors affecting dermal permeation/absorption include the thickness of the barrier (may vary depending on age, disease conditions and anatomical site), the lipid composition of the skin (Sato et al., 1991), as well as the density of hair follicles (permeation through the "shunt pathway"). Species differences regarding absorption rates of substances (Bartek et al., 1972) may be due, at least in part, to different levels of follicle density. Importantly, the presence of other substances in mixtures may change the bioavailability of a given substance. For instance, it has been shown that detergents and physical irritants work synergistically with respect to their effect on the impairment of the skin barrier, i.e. their combined effect is greater than the effect of both detergent and physical irritant applied on their own (Fluhr et al., 2005b,

2000c). The integrity of the dermal barrier (mainly maintained by the stratum corneum) is obviously of key importance for dermal irritation: an impaired stratum corneum renders the skin more permeable and irritants can more readily permeate into the living epidermal and eventual also the dermal layers increasing the susceptibility of such skin to irritation. The importance of a functioning barrier has been demonstrated in patients with atopic dermatitis (Cork et al., 2006): the irritant response to xenobiotics is increased in such patients and the strength of the irritant response correlates with the level of impairment of the stratum corneum (Al-Jaberi & Marks, 1984; Cowley & Farr, 1992). On a molecular level, the importance of the stratum corneum in the minimising susceptibility to irritant dermatitis has been underlined by a mutation in filaggrin, a protein that plays an role in stratum corneum homeostasis (Dereure, 2007; Houben E, 2007). Despite the importance of the stratum corneum, there are also irritants that do not or not to an significant extent disrupt the skin barrier (Fluhr et al., 2001; Snater et al., 1995). There is a variety of molecular mechanisms by which irritants may disrupt the barrier function of the skin. Anionic surfactants (such as Sodium Lauryl Sulfate) damage structural protein integrity (e.g. of keratin, the major component of the stratum corneum). This exposes water-binding sites in the proteins leading to hyperhydration of the stratum corneum and disruption of lipid bilayers (Fartasch, 1997; Ponc & Kempenaar J, 1995). In contrast organic solvents (benzene, toluol, acetone etc.) destroy the integrity of the skin barrier by extracting lipids from the stratum corneum, rendering it more permeable to a variety of substances (Fartasch, 1997).

Key event 2 - Metabolism

Metabolic transformation of xenobiotics (i.e. their bioactivation) may play a role in many toxic adverse effects, including skin irritation. Skin has innate metabolic capacity especially in the epidermal layers (Zhang et al., 2009; Steinsträsser & Merkle, 1995; Schaefer & Filaquier, 1992; Kao et al., 1985) and it is conceivable that substances may be transformed to more reactive metabolites by skin tissue. This has for instance been considered in the context of transepidermal drug delivery (discussed in Griesinger & Zuang, 2012; reviews: Zhang et al., 2009; Steinsträsser & Merkle, 1995).

Key event 3 - Tissue trauma triggered by chemical reactivity and other physicochemical interactions with molecular targets (biomolecules):

Tissue trauma, i.e. the impairment of the physiological functioning of cells (e.g. keratinocytes) is a universal trigger for inflammatory processes, irrespective of the tissue. Inflammation is the body's response to tissue trauma resulting from physical and chemical stressors, including xenobiotic substances carried by pathogens. In case of chemically-induced irritation, chemicals that bind to proteins residing in the extracellular matrix or in cell plasma membranes, or that bind to other moieties of cell plasma membranes (e.g. sugars) may impair the viability of cells in the affected skin tissue. This may include reactions such as protein denaturation, change of protein folding, saponification of membrane lipids. Morphological changes as a consequence of tissue trauma clearly precede irritant responses in man (Willis et al, 1989). Oxidative stress may also play a role (Willis et al., 1998). Apart from these general mechanisms leading to inflammatory skin reactions, some chemicals may act through more specific mechanisms or lead to sensations of irritation (sensory irritation) through targeted mechanisms. Examples are lactic acid or capsaicin activating TRPV1 channels in nerve fibres (TRPV1 = transient receptor potential cation channel subfamily V member 1). Moreover, it is conceivable that sensory nerve fibres (in dermis and epidermis) form

part of a complex regulatory loop of the innate immune system of the skin. There is evidence that peptidergic nerve fibres are required for both the induction and effector stage of contact sensitivity in mice (Beresford et al., 2004). Due to the fact that both endpoints are based to a considerable extent on local inflammation, these findings may have significance also for skin irritation.

Key event 4 - Release of inflammatory mediators by impaired cells

In response to tissue injury or trauma by irritant chemicals (key event 3), a complex cascade of nonspecific events is initiated involving inflammatory mediators, i.e. chemokines and cytokines. In skin inflammation these include Interleukines (IL) (IL-1 alpha, IL-5, IL-6, IL-8, IL-13, IL-33 as well as Tumor Necrosis Factor-alpha (TNF-alpha) and Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF) (Stamatas et al., 2013). Release of inflammatory mediators may happen via the regular secretory pathway or through non-classical export routes (Prudovsky et al, 2003; Prudovsky et al, 2008) including simple leakage of cytosolic inflammatory mediators from damaged skin cells into the interstitium (the intercellular space). The inflammatory mediators activate inflammatory pathways in the affected tissue (key event 5). There appears to be a relationship between the dose as well as the chemical nature of the irritant and the cytokine response evoked (Hoefakker et al., 1995; Spiekstra et al., 2005). One of the key players in dermal irritation appears to be Interleukin 1 alpha. Disruption of the plasma membrane barrier leads to the release of interleukin 1 alpha into the interstitium (Spiekstra et al., 2005; Wood et al., 1996; Mizutani et al., 1991). Interleukines in general appear to play a key role in the regulation of the expression and release of other cytokines and chemokines involved in inflammatory reactions. Primary signals involved in human dermal inflammation leading to irritation appear to be IL-1alpha and TNF-alpha leading to the release of secondary chemokine signals including CCL20 and CXCL8 (Spiekstra et al., 2005). CCL20 and CXCL8 have pro-inflammatory effects by activating epidermal residential cells involved in irritation by attracting a variety of immune cells including immature dendritic cells. Another CCL, CCL21 has been shown to be unregulated in response to exposure to chemical and physical irritants (Eberhard et al., 2004). CCL21 attracts dendritic cells to the site of injury (Saeki, 1999).

Key event 5 – Activation of the innate immune system and development of an inflammatory reaction of the skin tissue

Inflammatory mediators have a wide variety of effects on the tissue: they increase the diameter and permeability of blood vessels, attract innate immune cells (e.g. mast cells) and migratory immune cells (neutrophils, monocytes, T-cells, B-cells) to the site of injury and trigger the migration of migratory immune cells through the endothelium (leukodiapedesis) into the tissue where they participate in antigen clearance and tissue repair/remodelling. Mast cells residing in the dermal layers may play a key role as receptors of inflammatory mediator signalling. Mast cells are most numerous in the vicinity of nerve fibres, blood vessels and skin appendages. Mast cell granules contain histamine, heparin and vasoactive agents that may be released upon stimulation of the cells and these mediators may contribute to irritant responses (Tharp, 1991; Wasserman, 1990; Wasserman, 1983). In addition to these actions, inflammatory mediators may stimulate nerve endings leading to itching, stinging and burning sensations. Overall, the local inflammatory response leads to the classical clinical symptoms of skin irritation: redness (rubor), warmth (calor), painful sensations (dolor) and also the swelling (oedema) of the tissue area affected. Redness and warmth of the skin tissue are

consequences of increased blood flow of the inflamed area. The swelling of the skin is due to increased permeability of the endothelial cells forming the walls of blood vessels.

Chapter 3. Table of References with Notes

<i>Traditional animal test – OECD TG</i>	
TG available from: http://www.oecd.org/env/testguidelines	
OECD (2002). Acute Dermal Irritation/Corrosion, OECD Guideline for the Testing of Chemicals No. 404, OECD	This Test Guideline describes the in vivo test procedure in laboratory rabbits (healthy young adult albinos). The test substance is applied to a patch of skin on a dorsal area of the animal's trunk. The area should be freed of fur a day before the test. Test substances are applied to a small skin area (ca 6 cm ²) for normally 4 hours. Two physiological responses to the local inflammatory reaction are observed at 60 minutes, 24 , 48 and 72 hours after patch removal and up to 14 days. The reactions are erythema (skin reddening) and oedema (skin swelling). If these are reversible within 14 days, the substance is considered an irritant. If these are not reversible and/or if there other reactions (necrosis, bleeding etc.) the substance is considered corrosive.
<i>Accepted alternative test methods – OECD TGs and publications</i>	
TG available from: http://www.oecd.org/env/testguidelines	
<i>Skin corrosion</i>	
OECD (2004). In Vitro Skin Corrosion: Transcutaneous Electrical Resistance (TER), OECD Guideline for the Testing of Chemicals No. 430, OECD, Paris.	<p>The test guideline is based on the rat transcutaneous electrical resistance (TER) test method for skin corrosion assessment. The method uses excised skin discs as a test system. Loss of electrical resistance following application of a test substance indicates that the substance impairs the proper barrier function (based on an integer stratum corneum) and is thus indicative of corrosive properties.</p> <p>Current update of TG 430 (to be adopted by OECD in 2013) The update concerns the definition and inclusion of Performance Standards (PS) as Annex 1 to the TG. These PS will support the ready assessment and validation of similar and modified TER-based test methods in accordance with the principles of Guidance Document No. 34. The list of</p>

	reference chemicals (RCs) suggested in the updated version of the TG is identical to the RCs suggested in TG431.
OECD (2004). In Vitro Skin Corrosion: Human Skin Model Test, OECD Guideline for the Testing of Chemicals No. 431, OECD, Paris.	<p>The test guideline is based on Reconstructed human Epidermis (RhE) technology. The assay measures impairment of cell viability in the artificial skin construct as a readout of the damage caused by corrosives to skin. The RhE-based test methods for skin corrosion testing are accepted to distinguish corrosive substances (Category 1) from substances considered non-corrosive or non-classified (TG431; EU test method R40.bis).</p> <p>Current update of TG431 (to be adopted by OECD in 2013)</p> <p>The UN GH system foresees three further optional corrosive subcategories (1A, 1B, 1C) which are fully implemented in the EU through EU CLP but which are currently not predicted by the in vitro models. Notably, one of these RhE test methods, the EpiSkin, had originally been validated by ECVAM for subcategorisation of corrosives into R35 and R34, according to the previous DSD system (ESAC statement 1998). These categories correspond currently to subcategories 1A (=R35) and a combination of 1B and 1C (=R34). Information on subcategorisation is considered particularly important for the transport of chemicals: the three hazard subcategories directly correspond to UN packaging groups I to III. These differ substantially with regard to the restrictions imposed on the packaging and transport of substances (e.g. packaging size, dimension and quality of packaging material, means of transport etc.). Therefore, information on subcategorisation and hence packing groups would be useful additional information for chemicals producers and their downstream users in view of economising on packaging needs of chemicals that can be demonstrated to fall in packing groups II and III.</p> <p>To address this need, the OECD expert group on skin corrosion and irritation agreed that, on the background of the successful original validation of the EpiSkin (Fentem et al., 1998) for partial subcategorisation (R35 and R34), the currently available suite of RhE test methods for corrosion testing should be further evaluated in view of their ability to subcategorise. One of the major driving forces was the <i>availability of new control experiments</i>, not available during original validation. These control experiments address possible false predictions associated mainly with chemicals (a) that directly reduce the viability dye used in the protocols (MTT) or (b) interfere with an accurate measurement of optical density of formazan, the MTT reduction product (e.g. colorants).</p>

	<p>The projects provided the following results:</p> <ul style="list-style-type: none"> • None of the methods was able to resolve all three subcategories. However the SkinEthic assay was able to reliably and accurately predict category 1A versus a combination of categories 1B and 1C versus non-classified substances. • The remaining test methods (EpiDerm and SkinEthic) showed a high rate of over-predictions of 1B/1C categories as 1A (ca 45% over-prediction rate). • While this is not a concern from a precautionary point of view, it is an unacceptably high rate of false positive predictions. ECVAM has therefore suggested considering 1A predictions from these test methods as unresolved Category 1 predictions. This has been taken up in the draft TG (up for adoption by WNT in 2013). Thus, these test methods are able to predict Category 1, a combination of categories 1B and 1C and non-classified substances. • A very useful outcome of this project was the definition of a suitable list of reference chemicals (n=30) which have been included in the updated Performance Standards of TG431 and also TG 430. <p>It should be noted that this project was not a validation study, i.e. testing was performed only in the test developers' laboratories under non-blinded conditions and without using a consistent set of chemicals for all evaluated test methods.</p>
<p>OECD (2006). In Vitro Membrane Barrier Test Method for Skin Corrosion, OECD Guideline for the Testing of Chemicals No. 435, OECD, Paris.</p>	<p>The Corrositex method recapitulates the stratum corneum barrier of the skin using a matrix of hydrated collagen matrix in a supporting filter membrane. Corrosive chemicals or mixtures to pass through, by diffusion and/or destruction/erosion, this biobarrier and elicit a colour change in the underlying liquid "Chemical Detection System" (CDS), composed of water and pH indicator dyes. The assay has several limitations: many non-corrosive chemicals and mixtures and some corrosive chemicals and mixtures may not qualify for testing. For instance, aqueous substances with a pH range from 4.5 to 8.5 do typically not qualify for testing in the Corrositex assay. This excludes a lot of potential corrosives. While the Corrositex test method is accepted by OECD, it has not been included into EU legislation (i.e. test method regulation).</p>
<p><i>Skin irritation</i></p>	

<p>OECD (2010). In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method No. 439, OECD, Paris.</p>	<p>The test guideline is based on Reconstructed human Epidermis (RhE) technology. The assay measures impairment of cell viability in the artificial skin construct as a readout of the cell damage and tissue trauma, the initial step in the inflammatory cascade leading to skin irritation. The test methods are accepted for predicting skin irritants (Category 2) versus non-classified substances. They are not able to predict the optional Category 3 of UN GHS (UN, 2013), not implemented in the EU (EU, 2008).</p> <p>Current update of TG 439 (to be adopted by OECD in 2013) Following a performance-based validation study of the Japanese LabCyte EPI-MODEL (manufactured by J-Tec, Ltd), the Japanese Centre for the Validation of Alternative Methods had requested an evaluation of this study by the OECD expert group. Following positive evaluation of the (amended) study by the OECD expert group, the Working Group of the National Coordinators of the Test Guidelines Programme (WNT) decided in 2012 to include this assay in TG 439. On request of the OECD Test Guidelines Programme, ECVAM acted as lead organisation and prepared, in collaboration with JaCVAM, the necessary update of TG439, taking the opportunity to amend the test guideline also with regard to a) the clarity of presentation of the underlying validation evidence for each of the methods and b) an overview on the main communalities and differences of the various test methods' protocols included in TG 439. The latter should support test users with regard to selecting the test methods that meets their specific requirements. Such a comparison table adds moreover to the transparency of the test methods included in the TG and may become a standard element in Performance-Based Test Guidelines (PBTGs). The updated version is currently (July 2013) scheduled for adoption by OECD within 2013.</p>
<p>Griesinger C, Barroso J, Zuang V, Cole T, Genschow E & Liebsch M (2010). Explanatory background document to the OECD Test Guideline on in vitro skin irritation testing. OECD Series on Testing and Assessment, No. 137, OECD, Paris.</p> <p><u>An updated version is made available from the EURL ECVAM webpage:</u> "Explanatory Background Document to the OECD</p>	<p>This background document compiles comprehensive information on the ECVAM validation study and other validation studies relating to the RhE test methods of TG439 for skin irritation testing. It provides a summary of the scientific work leading the validated RhE methodology, the adaptations of the performance standards reference chemicals and accuracy values necessary as a consequence of the change from the former EU DSD-based classification system to the UN GHS-based EU CLP system. The document provides a detailed description (including structural formula) of the test chemicals used for validation. Finally, also the results from the preceding optimisation studies are presented.</p>

<p>Draft Test Guideline on in vitro Skin Irritation Testing. 1. Performance of Test Methods under UN-GHS; 2. Update of Performance Standards; 3. Comprehensive Data Compilation pursuant to the Validation Process and the preceding Optimisation Studies." http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/validation-regulatory-acceptance/topical-toxicity/skin-irritation</p>	
<p><i>Traditional animal test – Scientific publications</i></p>	
<p>Draize JH, Woodard G & Calvery HO (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J. Pharmacol. and Exp. Therapeutics. 82, 377–390.</p>	<p>The paper suggests the use of rabbits for testing irritation and toxicity (corrosion) of substances, applied to the skin and mucous membranes (including eyes) of the animals.</p>
<p>Weil CS, Scala RA (1971). Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol Appl Pharmacol. Jun;19(2):276-360.</p>	<p>This paper analyses the intra- and interlaboratory variability of Draize eye and skin irritation testing data on the basis of 10 substances distributed to 24 laboratories. The study finds within- and between laboratory variability.</p>
<p>ECETOC (1995). Skin Irritation and Corrosion Reference Chemicals Data Bank. ECETOC technical report No. 66. Brussels, Belgium.</p>	<p>This Reference Chemicals Data Bank provides animal testing data of high quality (i.e. performed under GLP). Analysis of the variability of these data shows high variability. An analysis is presented in Griesinger C et al., Explanatory background document to the OECD draft Test Guideline on in vitro skin irritation testing. OECD 2010. OECD Series on Testing and Assessment, No. 137, OECD, Paris.</p>
<p>Phillips L 2nd, Steinberg M, Maibach HI, Akers WA (1972). A comparison of rabbit and human skin response to certain irritants. Toxicol Appl Pharmacol. 21(3):369-82.</p>	<p>An early comparative analysis of animal Draize irritation versus human data which finds differences in the irritant responses.</p>
<p>Jírová D, Basketter D, Liebsch M, Bendová H, Kejlová K, Marriott M, Kandárová H (2010). Comparison of human skin irritation patch test data with in vitro skin irritation assays and animal data. Contact Dermatitis 62(2):109-16.</p>	<p>The paper presents an interesting comparison between (a) human skin irritation data with (b) in vivo irritation data and (c) data from in vitro RhE skin irritation tests (using EpiDerm and EpiSkin). Background to this analysis is the high variability of animal reference data and the issues with regard to relevance for the species of interest, the human. Using the 4-hour human patch test (HPT) the authors examined 16 chemicals</p>

	whose EU classification of skin irritancy is known to be borderline (i.e. at the border between Category 2 and No Category), or for which in vitro methods have been reporting conflicting results. Of the 16 chemicals classified as irritants in the rabbit, five substances were found to be irritant to human skin. Concordance of the rabbit test with the 4-hr HPT was only 56%, whereas concordance of human epidermis models with human data was 76% (EpiDerm) and 70% (EpiSkin). The study confirms that the rabbit model is overpredicting (which from a precautionary point of view needn't be a worry), but which has drawbacks for industry. Furthermore, the study shows that in vitro methods based on human cells may have a higher relevance with respect to human effects when comparing the in vitro data with data from a human standardised irritation test (HPT).
Basketter DA, York M, McFadden JP, Robinson MK (2004). Determination of skin irritation potential in the human 4-h patch test. <i>Contact Dermatitis</i> . 51(1):1-4.	The paper presents a substantial database on the skin irritation potential of some 65 substances tested in a standard human 4-h patch test. These provide a high quality dataset for the development, evaluation and validation of <i>in vitro</i> or <i>in silico</i> alternatives. This allows avoiding the optimisation of alternative methods against non-human reference data from the rabbit.
Young JR, How MJ, Walker AP, Worth WMH (1988). Classification as Corrosive or Irritant to Skin of Preparations Containing Acidic or Alkaline Substance Without Testing on Animals. <i>Toxicol. In Vitro</i> 2, 19-26.	The paper suggest that materials that for substances with strong acidity or alkalinity, animal testing is not necessary. This was taken up in the testing strategy described in the supplement of TG 404.
<i>Accepted alternative test methods – Scientific publications</i>	
<i>Skin Corrosion</i>	
Barratt MD, Brantom PG, Fentem JH, Gerner I, Walker AP, Worth AP (1998) The ECVAM international validation study for in vitro tests for skin corrosivity. 1. Selection and distribution of the test chemicals. <i>Toxicol. In Vitro</i> 12:471-482.	These papers describe the ECVAM validation study on skin corrosion assays, involving the following test methods TER (later TG430), Skin(2TM) ZK1350 test method, Corrositex (later TG435) and EpiSkin test methods (later TG431). While the paper by Barratt et al., focuses on the test selection of appropriate test substances, the publication by Fentem et al, provides the results of the validation study.
Fentem JH, Archer GE, Balls M, Botham PA, Curren	

RD, Earl LK, Esdaile DJ, Holzhütter HG, Liebsch M. (1998) The ECVAM International Validation Study on In Vitro Tests for Skin Corrosivity. 2. Results and Evaluation by the Management Team. Toxicol In Vitro 12(4):483-524.	
Kandárová H, Liebsch M, Spielmann H, Genschow E, Schmidt E, Traue D, Guest R, Whittingham A, Warren N, Gamer AO, Remmele M, Kaufmann T, Wittmer E, De Wever B, Rosdy M. (2006). Assessment of the human epidermis model SkinEthic RHE for in vitro skin corrosion testing of chemicals according to new OECD TG 431. Toxicol In Vitro 20(5):547-59.	The paper summarises the SkinEthic validation study based on the TG431 Performance Standards for RhE based skin corrosion tests involving assessment of 12 reference chemicals.
Tornier C, Roquet M, Fraissinette AB (2010). Adaptation of the validated SkinEthic™ Reconstructed Human Epidermis (RHE) skin corrosion test method to 0.5 cm ² tissue sample. Toxicol. In Vitro 24, 1379-1385.	The publication describes the adaptation of the SkinEthic validated method to a smaller tissue insert size and provides information on the performance of this slightly modified method.
Hoffmann J, Heisler E, Karpinski S, Losse J, Thomas D, Siefken W, Ahr HJ, Vohr HW, Fuchs HW (2005). Epidermal-skin-test 1,000 (EST-1000) - a new reconstructed epidermis for in vitro skin corrosivity testing. Toxicol In Vitro 19(7):925-9.	The paper describes the EST-1000 validation study based on the TG431 Performance Standards for RhE based skin corrosion tests involving assessment of 12 reference chemicals. From 2012 on, the EST-1000 model has been marketed under the trade name "EpiCS". The model's technical specifications have not changed.
<i>Skin Irritation</i>	
Fentem JH, Briggs D, Chesné C, Elliot GR, Harbell JW, Heylings JR, Portes P, Roguet R, van de Sandt JJM and Botham P (2001). A prevalidation study on in vitro tests for acute skin irritation, Results and evaluation by the Management Team. Toxicol. in Vitro 15, 57-93.	The paper describes the ECVAM prevalidation study (1999-2000) on a selection of in vitro / ex vivo skin irritation tests. These were: <ul style="list-style-type: none"> • the EpiDerm assay, • the EPISKIN assay, • the PREDISKIN assay, • the non-perfused pig ear method, • the mouse skin integrity function test (SIFT).

	Neither PREDISKIN nor the pig ear method performed sufficiently well to proceed to the final stage of the Prevalidation study. The SIFT assay needed further refinement work and did not proceed to the last stage. Only the two RhE-based methods (EpiDerm and EpiSkin) successfully underwent all phases of the prevalidation study. However, none of the methods was found ready to enter into a formal validation study and improvement of the protocols and prediction models was recommended.
Zuang V, Balls M, Botham PA, Coquette A, Corsini E, Curren RD, Elliot GR, Fentem JH, Heylings JR, Liebsch M, Medina J, Roguet R, van De Sandt JJM, Wiemann C and Worth A (2002). Follow-up to the ECVAM prevalidation study on in vitro tests for acute skin irritation, The European Centre for the Validation of Alternative Methods Skin Irritation Task Force report 2. ATLA 30, 109-129.	The paper summarised the outcome of an ECVAM task force meeting on skin irritation during which the results of the Prevalidation study (Fentem et al., 2001) were discussed and during which decisions were taken with regard to the test methods that were found sufficiently well developed to enter formal validation.
Kandárová H, Liebsch M, Genschow E, Gerner I, Traue D, Slawik B and Spielmann H (2004). Optimisation of the EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests. ALTEX 21, 107–114. Kandárová H, Liebsch M, Gerner I, Schmidt E, Genschow E, Traue D and Spielmann H (2005). The EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests – An assessment of the performance of the optimised test. ATLA 33, 351-367.	These publications summarise the test method optimisation work of the EpiDerm model prior to its validation by ECVAM.
Cotovio J, Grandidier MH, Portes P, Roguet R and Rubinsteen G (2005). The in vitro acute skin irritation of chemicals: optimisation of the EPISKIN prediction model within the framework of the ECVAM validation process. ATLA 33, 329-349.	This publication summarises the test method optimisation work of the EpiSkin model prior to its validation by ECVAM.
Spielmann H, Hoffmann S, Liebsch M, Botham P,	This publication summarises the findings of the ECVAM validation study on the RhE test

<p>Fentem JH, Eskes C, Roguet R, Cotovio J, Cole T, Worth A, Heylings J, Jones P, Robles C, Kandárová H, Gamer A, Remmele M, Curren R, Raabe H, Cockshott A, Gerner I, Zuang V (2007). The ECVAM international validation study on in vitro tests for acute skin irritation: report on the validity of the EPISKIN and EpiDerm assays and on the Skin Integrity Function Test. ATLA 35(6):559-601.</p>	<p>methods for skin irritation testing (EpiDerm and EpiSkin).</p>
<p>Eskes C, Cole T, Hoffmann S, Worth A, Cockshott A, Gerner I and Zuang V (2007). The ECVAM international validation study on in vitro tests for acute skin irritation: selection of test chemicals. ATLA 35, 603-619.</p>	<p>The paper provides a detailed summary of the test chemicals used (n=58) in the ECVAM skin irritation validation study and provides information on the chemical selection process using a set of pre-defined criteria.</p>
<p>Katoh M and Hata K (2011). Refinement of LabCyte EPI-MODEL24 skin Irritation test method for adaptation to the requirements of OECD test guideline 439. AATEX, 16, 111-122</p> <p>Katoh M, Hamajima F, Ogasawara T and Hata K (2009). Assessment of human epidermal model LabCyte EPI-MODEL for in vitro skin irritation testing according to European centre for the validation of alternative methods (ECVAM)-Validated Protocol. J Toxicol Sci, 34, 327-334</p> <p>Kojima H, Ando Y, Idehara K, Katoh M, Kosaka T, Miyaoka E, Shinoda S, Suzuki T, Yamaguchi Y, Yoshimura I, Yuasa A, Watanabe Y and Omori T (2012). Validation Study of the In Vitro Skin Irritation Test with the LabCyte EPI-MODEL24. ATLA 40, 33-50.</p>	<p>These papers summarise the optimisation and validation of the Japanese LabCyte model. The validation study was peer reviewed by OECD in 2011 and the test method adopted for TG439 in 2012. The updated TG439 is expected to be formally adopted by OECD in 2013.</p>

Regulatory requirements across sectors

European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L396: 1-849.

REACH/Chemicals: Information requirements depend on the production volume of the chemical concerned. While Annex VII foresees only *in vitro* testing, Annex VIII foresees the traditional animal test.

Annex VII stipulates the information requirements for substances manufactured or imported in the EU in quantities of one ton or more per year (lowest tonnage level); these include skin irritation and corrosion. The assessment of these effects should follow four consecutive steps: (1) assessment of available human and animal data; (2) assessment of acid or alkaline reserve; (3) *in vitro* study for skin corrosion; (4) *in vitro* study for skin irritation.

Annex VIII stipulates the information requirements for substances manufactured or imported in the EU in quantities of 10 tonnes or more per year. It foresees the assessment of skin irritation using the *in vivo* test.

Furthermore, with Annex XI, REACH allows adaptation of the standard testing regime outlined in Annexes VII to X and thus, in principle, allows the adaptation of the *in vivo* testing requirement set out in Annex VIII (Annex VI in any case foresees only *in vitro* testing) beyond the specific rules laid out in Column 2 of Annex VIII. Possible ways of adapting the standard regime are Annex XI 1.2 (weight of evidence) and Annex XI 1.4 (*in vitro* methods).

European Commission (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353: 1-1355.

Classification, Labelling and Packaging: The CLP Regulation itself does not stipulate information requirements, but lays out the rules for Classification and Labelling. It complements in particular the REACH legislation: the aim of human health hazard assessment under REACH (Annex I, section 1.) is after all the determination of the classification and labelling of a substance in agreement with the relevant rules (formerly Directive 67/548/EEC, replaced by CLP Regulation).

The EU CLP Regulation implements the UN Globally Harmonized System (UN GHS) for classification and labelling (C&L). The C&L categories used are based on visually

	<p>observable effects in rabbit skin following exposure (Draize skin corrosion and skin irritation test, EU test method B.4; OECD Test Guideline 404).</p> <p>Corrosive substances are labelled 'Category 1'. This category contains three further optional subcategories which correspond to the UN Packing Groups I, II and III for the transport of goods. The subcategories are implemented in the EU. They differ with regard to the exposure times required to elicit skin corrosion in the rabbit and are referred to as 1A ("strong corrosives"), 1B ("moderate corrosives") and 1C ("mild corrosives").</p> <p>Irritant substances are labelled 'Category 2'. While UN GHS allows furthermore the use of an additional "opt-in" category to classify mild irritants ('Category 3'), this optional category is not implemented in the EU and hence, C&L of skin irritants in the EU is follows a dichotomous categorisation: Category 2 versus non-classified.</p>
<p>European Commission (2009a). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309: 1-47.</p>	<p>Plant Protection Products (PPPs): The PPP Regulation requests information on skin irritation and corrosion (section 7.1.4; note that the title of the section refers only to irritation, while the text refers to both, irritation and corrosion). The Regulation stipulates that, before carrying out in vivo studies for corrosion/irritation of the PPP, a Weight of Evidence (WoE) analysis should be carried out. This strategy is similar to that outlined in the supplement to OECD TG 439: If the data evaluated by WoE is insufficient to reach conclusions on corrosion/irritation potential of the PPP, a testing strategy should be carried out using (a) a validate in vitro method for skin corrosion; (b) assessment of irritation using a validated in vitro test method (such as human reconstructed skin models); (c) an initial in vivo study using one animal and (d) confirmatory testing using one or two animals.</p> <p>Information on corrosion and irritation may also be derived from dermal toxicity studies (section 7.1.2), however such studies are required only on a case by case basis under the PPP Regulation.</p>
<p>European Commission (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342: 59-209.</p>	<p>Cosmetics: The Cosmetics Regulation requests a characterisation of the "toxicological profile of substance contained in the cosmetic product for all relevant toxicological endpoints. A particular focus on local toxicity evaluation (skin and eye irritation...)" (Annex I, point 8 of Cosmetics Regulation). Clearly the evaluation of the corrosion and irritation potential of substances voluntarily and repeatedly applied to skin is obviously of great relevance.</p>

<p>European Union Regulation (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167: 1-116.</p>	<p><i>Biocides:</i> All biocidal products and active substances marketed in the UE independent of tonnage level. Information requirement for skin irritation and skin corrosion follows a strategic approach using in vitro methods for skin corrosion and irritation and, only in case of specific limitations of these in vitro methods, in vivo testing should be considered. (ECHA guidance on Biocides Regulation, section 8.1).</p>
<p><i>Mechanistic understanding</i></p>	
<p>Elsner P (1994). Irritant dermatitis in the workplace. <i>Dermatol Clin.</i> 12(3):461-7</p>	<p>The article reviews pathogenesis, epidemiology, diagnosis, and treatment of irritant dermatitis at the workplace. Irritant contact dermatitis is the most frequent occupational skin disease and results in considerable morbidity and economic losses. The importance of integrated preventive measures (that require robust knowledge about the irritancy potential of substances) are suggested which should be implemented by companies and employees to achieve optimal efficacy.</p>
<p>Wigger-Alberti W, Elsner P (2000). Contact Dermatitis due to Irritation, p99-110. In <i>Handbook of Occupational Dermatology</i>. Berlin, Springer.</p>	<p>The paper outlines the clinical relevance of Irritant Contact Dermatitis (ICD). Skin diseases are the second most frequent occupation disease (after musculoskeletal disorders, e.g. "back pain"). Most dermatoses (diseases of skin) are cases of contact dermatitis and ICD is more frequent than Allergic Contact Dermatitis (ACD) resulting from skin sensitisation. In contrast to ACD, ICD is defined as a result of primarily nonspecific damage (trauma) to the skin and hence does not represent a precise entity in clinical terms but a spectrum of disease patterns. The authors argue that the perception of ICD as more trivial than the more intellectually appealing problem of allergic sensitisation (ACD) is changing. The review stresses the eminent importance of reliable data on skin corrosion and irritation potential of substances.</p>
<p>Willis CM, Stephens CJ, Wilkinson JD (1989). Epidermal damage induced by irritants in man: a light and electron microscopic study. <i>J Invest Dermatol.</i> 93(5):695-9.</p>	<p>This paper is one of the very few thorough contributions to the study of the pathogenesis of Irritant Contact Dermatitis by studying histopathological damages in human skin resulting from exposure to a series of structurally unrelated irritants (i.e. reflecting a diverse chemistry in view of determining possible differences in histopathology). Human volunteers were patch-tested with nonanoic acid, sodium lauryl sulphate, dithranol, benzalkonium chloride, croton oil, and propylene glycol, which produced generally mild to moderate responses. Biopsy specimens were excised 48 h after topical application of the irritants and examined by light and electron microscopy. Spongiosis and infiltration of predominantly mononuclear cells were observed in the epidermis.</p>

	<p>Several irritants induced characteristic patterns of keratinocyte damage, others caused morphologic changes indicative of disturbances in keratinocyte metabolism and differentiation. The results suggest that there is a diversity and specificity in the histopathology of irritant contact dermatitis, reflecting the different ways in which chemicals may interact with components of the skin. However, tissue trauma resulting from severe impairment of cell viability appears a common consequence irrespective of the irritant. The authors have published four other papers on ultrastructural effects following irritant exposure in humans. These are not presented in greater detail, but listed below:</p> <p>Willis CM, Stephens CJ, Wilkinson JD (1991). Selective expression of immune-associated surface antigens by keratinocytes in irritant contact dermatitis. <i>J Invest Dermatol.</i> 96(4):505-11.</p> <p>Willis CM, Stephens CJ, Wilkinson JD (1990). Differential effects of structurally unrelated chemical irritants on the density and morphology of epidermal CD1+ cells. <i>J Invest Dermatol.</i> 95(6):711-6.</p> <p>Willis CM, Stephens CJ, Wilkinson JD (1992). Differential effects of structurally unrelaed chemical irritants on the density of proliferating keratinocytes in 48 h patch test reactions. <i>J Invest Dermatol.</i> 99(4):449-53.</p> <p>Willis CM, Stephens CJ, Wilkinson JD (1993). Differential patterns of epidermal leukocyte infiltration in patch test reactions to structurally unrelated chemical irritants. <i>J Invest Dermatol.</i> 101(3):364-70.</p>
<p>Willis CM, Reiche L, Wilkinson JD (1998). Immunocytochemical demonstration of reduced Cu,Zn-superoxide dismutase levels following topical application of dithranol and sodium lauryl sulphate: an indication of the role of oxidative stress in acute irritant contact dermatitis. <i>Eur J Dermatol.</i> 8(1):8-12.</p>	<p>This paper demonstrates the involvement of oxidative stress in the genesis of chemically induced skin inflammation / skin irritation leading to Irritant Contact Dermatitis (ICD).</p>
<p>Willis CM, Young E, Brandon DR, Wilkinson JD (1986). Immunopathological and ultrastructural</p>	<p>This paper compares the histopathological features of Allergic Contact Dermatitis (ACD) with those of Irritant Contact Dermatitis (ICD) in a group of 17 patients that each patient</p>

findings in human allergic and irritant contact dermatitis. Br J Dermatol. 115(3):305-16.	received simultaneous patch tests of a known allergen and a standardized irritant (benzalkonium chloride). Cellular changes occurring between 3 h and 7 days after patch test application were studied by light and electron microscopy and immunocytochemistry. Interestingly the paper did not find any differences between the induced ACD and the ICD, neither in the responding cell types nor the sequence of cellular events. In both cases infiltration mainly by T lymphocyte occurred (no polymorphonuclear leukocyte involvement). Apposition of Langerhans cells to lymphocytes in the epidermis was seen in both types of response. While there was considerable variability in the intensity of reaction to irritant and allergen occurred within individuals, no statistically significant difference between the intensity of the reactions to the irritant and the allergen was observed in the patient group tested. The paper demonstrates that the early inflammatory response mediated mainly by the innate immune system of the skin tissue is the same for ACD and ICD.
Stamatas GN, Morello AP Mays DA (2013). Early Inflammatory Processes in the Skin. Curr Mol Med. Feb 27	This review focuses on immunologic activity occurring in the absence of any visual inflammatory cues. The authors discuss the importance of subclinical inflammation in human skin and its relevance to innate immune surveillance under physiologic conditions.
Fluhr JW, Darlenski R, Angelova-Fischer I, Tsankov N, Basketter D (2008) Skin irritation and sensitization: mechanisms and new approaches for risk assessment. 1. Skin irritation. Skin Pharmacol Physiol. 21(3):124-135.	This review outlines advanced techniques for studying irritation as well as an overview of key inflammatory mediators involved in skin irritation.
<i>Status of non-standard methods / Status of new in-vitro based standard methods</i>	
Bouvier d'Yvoire M, Bremer S, Casati S, Ceridono M, Coecke S, Corvi R, Eskes C, Gribaldo L, Griesinger C, Knaut H, Linge JP, Roi A, Zuang V (2012). ECVAM and New Technologies for Toxicity Testing in New Technologies for Toxicity Testing, eds Balls M, Combes RD, Bhogal N. Adv Exp Med Biol. 745:154-80.	This book chapter contains a section on skin corrosion/irritation where new developments are presented in detail.
Poumay Y, Dupont F, Marcoux S, Leclercq-Smekens M, Hérin M, Coquette A (2004). A simple	This paper describes a method for reconstructing human epidermis in vitro. The method was developed by university and is devoid of any proprietary elements. The protocol of Poumay and co-authors has later been taken up by Henkel, Germany, who is developing

<p>reconstructed human epidermis: preparation of the culture model and utilization in in vitro studies. Arch Dermatol Res. 296(5):203-11.</p>	<p>it further in view of marketing this "open source" model for standardised use.</p>
<p><i>Integrated Testing Strategies</i></p>	
<p>OECD (2004). Acute Dermal Irritation/Corrosion, OECD Guideline for the Testing of Chemicals No. 404, OECD</p>	<p>The TG contains a supplement outlining a simple testing strategy. The supplement recommends that a weight-of-the-evidence analysis may be used to evaluate existing information regarding the skin irritation and corrosion of substances to determine whether additional studies, other than in vivo dermal studies, should be performed to help characterise such potential. Although this sequential testing strategy is not an integral part of Test Guideline 404, it expresses the recommended approach for the determination of skin irritation/corrosion characteristics. The strategy provides an approach for the evaluation of existing data on the skin irritation/corrosion properties of test substances and a tiered approach for the generation of relevant data on substances for which additional studies are needed, or for which no studies have been performed. It also recommends the performance of validated and accepted <i>in vitro</i> or <i>ex vivo</i> tests for skin corrosion/irritation under specific circumstances.</p>
<p>ECHA (2012). Guidance on information requirements and chemical safety assessment. Chapter R.7a: Endpoint specific guidance. In "Guidance for the implementation of REACH". Version 2.0. November 2012.: http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf</p>	<p>This document provides guidance on information requirements including skin corrosion and irritation. It contains the testing strategy developed during the REACH implementation project.</p>
<p>Hoffmann S, Saliner AG, Patlewicz G, Eskes C, Zuang V & Worth AP (2008). A feasibility study developing an integrated testing strategy assessing skin irritation potential of chemicals. Toxicology Letters 180: 9–20.</p>	<p>This paper describes a feasibility study investigating how a combination of in silico, in vitro, and in vivo information could be applied in the assessment of skin irritation hazard. Therefore, a database of 100 existing and new chemicals was compiled. A number of strategies, both animal-free and inclusive of animal data were constructed and subsequently evaluated considering predictive capacities, severity of misclassifications and testing costs. Comparison of constructed ITS based on these assessment parameters identified best performing strategies for chemical classification. However, defining the in vivo test as the reference test limited the evaluation of the ITS inclusive of animal data.</p>

	This study demonstrated that ITS can be constructed, evaluated and compared in a systematic fashion. To promote ITS, further guidance on construction and multi-parameter evaluation need to be developed.
Worth AP, Fentem JH, Balls M, Botham PA, Curren D, Earl LK, Esdaile DJ & Liebsch M (1998). An Evaluation of the Proposed OECD Testing Strategy for Skin Corrosion. 709–720.	The widespread concern over the severity of the Draize rabbit test for assessing skin irritation and corrosion led to the proposal of a stepwise testing strategy at an OECD workshop in January 1996. Subsequently, the proposed testing strategy was adopted, with minor modifications, by the OECD Advisory Group on Harmonization of Classification and Labelling. This article reports an evaluation of the proposed OECD testing strategy as it relates to the classification of skin corrosives. By using a set of 60 chemicals, an assessment was made of the effect of applying three steps in the strategy, taken both individually and in sequence. The results indicate that chemicals can be classified as corrosive (C) or non-corrosive (NC) with sufficient reliability by the sequential application of three alternative methods, i.e., structure-activity relationships (where available), pH measurements, and a single in vitro method (either the rat skin transcutaneous electrical resistance (TER) assay or the EPISKINTM assay). It is concluded that the proposed OECD strategy for skin corrosion can be simplified without compromising its predictivity. For example, it does not appear necessary to measure acid/alkali reserve (buffering capacity) in addition to pH for the classification of pure chemicals.
<i>QSARs and Expert Systems</i>	
Barratt MD (1996a). Quantitative structure-activity relationships (QSARs) for skin corrosivity of organic acids, bases and phenols: Principal components and neural network analysis of extended datasets. Toxicology in Vitro 10: 85–94.	Quantitative structure-activity relationships (QSARs) relating skin corrosivity data of organic acids, bases and phenols to their log(octanol/water partition coefficient), molecular volume, melting point and pK(a). have been extended to substantially larger datasets. In addition to principal components analysis, as used in earlier work, the datasets have also been analysed using neural networks. Plots of the first two principal components of the four independent variables, which broadly model skin permeability and cytotoxicity, for each of the extended datasets confirmed that the analysis was able to discriminate well between corrosive and non-corrosive chemicals. Neural networks using the same parameters as inputs, were trained to an output in the range 0.0 to 1.0, with non-corrosive chemicals being assigned the value 0 and corrosive chemicals the value 1. As well as yielding classification predictions in agreement with those in the training sets,

	<p>predicted outputs in the 0 to 1 range gave a useful indication of the confidence of the predicted classification. These QSARs are useful (a) for the prediction of the skin corrosivity potentials of new or untested chemicals and (b) for determining the confidence of predictions in regions of 'biological uncertainty' which exist at the classification threshold between corrosive and non-corrosive chemicals.</p>
<p>Barratt MD (1996b). Quantitative structure-activity relationships for skin irritation and corrosivity of neutral and electrophilic organic chemicals. <i>Toxicology in Vitro</i> 10: 247–256.</p>	<p>Quantitative structure-activity relationships (QSARs) have been derived by relating skin irritation and corrosivity data of neutral and electrophilic organic chemicals to their log(octanol/water partition coefficient) (logP), molecular volume, dipole moment and 1/molecular weight. Datasets were analysed using stepwise regression, discriminant and principal components analysis. Discriminant analysis between irritant and non-irritant neutral and electrophilic organic chemicals using the above parameters, which broadly model skin permeability (logP and molecular volume), 'reactivity' (dipole moment) and 1/molecular weight to compensate for the fact that skin irritation/corrosivity testing is carried out using a fixed mass or volume of chemical, was found to discriminate well for only 73.1% of the dataset (67.3% cross-validated). The poor discrimination at the irritant/non-irritant classification boundary is attributed largely to biological variability. Stepwise regression analysis of the Primary Irritation Index (PII) for the same dataset showed a poor correlation ($r(2) = 0.422$; cross-validated $r(2) = 0.201$) with a positive dependence on logP and dipole moment and a negative dependence on molecular volume; 1/molecular weight was not a significant variable. While this QSAR for PII has little value as a predictive model, mainly because of the large biological variability evident in PII values, it is useful in confirming the putative model for skin irritation. Discriminant analysis using logP, molecular volume and dipole moment, was able to discriminate reasonably well (92.9% well-classified; 92.9% cross-validated) between corrosive and non-corrosive electrophiles. A plot of the first two principal components of the same parameters showed a clear demarcation between corrosive and non-corrosive electrophiles. In contrast to the QSARs for skin irritation, increasing skin corrosivity was found to correlate with decreasing molecular volume, with increasing dipole moment, and with decreasing logP. The predominant parameter in determining the skin corrosivity of electrophilic organic chemicals appears to be the molar dose at which they are tested; this arises because skin corrosivity testing is conducted using a fixed mass or volume of chemical. A stepwise approach to the skin corrosivity/irritation classification of neutral and electrophilic organic chemicals is outlined. The derived QSARs should be useful for</p>

	<p>the prediction of the skin corrosivity potential of new or untested electrophiles. (Non-electrophilic neutral organic chemicals, as a category, do not generally appear to be corrosive.) Discrimination between some non-irritant and irritant neutral and electrophilic organic chemicals using these techniques is also possible. For a large number of chemicals whose irritation potentials lie in a fairly broad band around the irritant/non-irritant classification boundary, no firm prediction of classification is possible.</p>
<p>Gallegos Saliner A, Patlewicz G & Worth AP (2006). Review of Literature-Based Models for Skin and Eye Irritation and Corrosion. JRC Report EUR 22320 EN. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/QSAR_Review_Irritation.pdf</p>	<p>This report reviews the state-of-the-art of <i>in silico</i> and <i>in vitro</i> methods for assessing dermal and ocular irritation and corrosion. Following a general introduction, the current EU legislation for the classification and labelling of chemicals causing irritation and corrosivity is summarized. Then currently available non-animal approaches are reviewed. The main alternative approaches to assess acute local toxic effects are: a) <i>in silico</i> approaches, including SARs, QSARs and expert systems integrating multiple approaches; and b) <i>in vitro</i> test methods. In this review, emphasis is placed on literature-based (Q)SAR models for skin and eye irritation and corrosion as well as computer-based expert systems.</p>
<p>Gallegos Saliner A, Patlewicz G & Worth AP (2008). A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. <i>QSAR & Combinatorial Science</i> 27: 49-59.</p>	<p>This paper reviews the state-of-the-art of <i>in silico</i> methods for assessing dermal and ocular irritation and corrosion. It is based on an in-depth review performed by the European Chemicals Bureau of the European Commission: Joint Research Centre in support of the development of technical guidance for the implementation of the REACH legislation, and is one of a series of minireviews in this journal. The most widely used <i>in silico</i> approaches are classified into methods to assess (1) skin irritation, (2) skin corrosion and (3) eye irritation. In this review, emphasis is placed on literature-based (Q)SAR models.</p>
<p>Gerner I, Schlegel K, Walker JD & Hulzebos E (2004). Use of Physicochemical Property Limits to Develop Rules for Identifying Chemical Substances with no Skin Irritation or Corrosion Potential. <i>QSAR Combinatorial Science</i> 23:726-733.</p>	<p>This paper describes limit values for specific physicochemical properties that are appropriate for identifying chemical substances that have no skin irritation or corrosion potential. These physicochemical properties include melting point, molecular weight, octanol-water partition coefficient, surface tension, vapour pressure, aqueous solubility and lipid solubility. Based on analyses of 1833 chemicals, physicochemical properties for limits were defined to determine that when a chemical's physicochemical properties were either greater or less than these limits that these chemicals would have no skin irritation or corrosion potential. To facilitate classification and labeling, the application domains of these limits were constructed to correspond with the European Union's risk phrases for chemicals classified for skin irritation/corrosion, viz., R 34, R35 or R38. This is the</p>

	<p>second paper of four companion papers. The first paper discussed mechanisms that can lead to significant skin irritation or corrosion after acute exposures to chemicals. The third paper described the application of structural alerts to identify chemical substances with skin irritation or corrosion potential. The fourth paper described the Skin Irritation Corrosion Rules Estimation Tool (SICRET), a tool that allows non-(Q)SAR experts to identify chemical substances with skin irritation or corrosion potential based on physicochemical property limits and structural alerts.</p>
<p>Golla S, Madihally S, Robinson RL & Gasem KAM (2009). Quantitative structure-property relationships modeling of skin irritation. <i>Toxicology in vitro</i>: 23:176–84.</p>	<p>Interest in developing procedures for estimating skin irritation potential of chemicals has been increasing as a result of concerns regarding animal welfare and costs involved in experimental irritation studies. In response to these concerns, a number of expert systems and quantitative structure-activity relationship (QSAR) models have been proposed for predicting the skin irritation potential of compounds. However, these models require as input independent estimates of several physicochemical properties. Hence, to predict skin irritation potential using these models often requires additional models capable of estimating the physicochemical properties of diverse structures; a requirement that most literature QSARs fail to meet. In the work reported here, we developed a skin irritation QSPR model based on rabbit Draize test data for 186 compounds, which included chemicals from diverse molecular classes. The effectiveness of using a combination of traditional, functional group and structural descriptors has been studied. Our non-linear QSPR model is capable of predicting the skin irritation potential of chemical compounds with an R² of 0.78. Further, the final set of descriptors used to model skin irritation was analyzed to elucidate the effects of molecular size, reactivity and skin penetration on skin irritation.</p>
<p>Hayashi M et al. (1999). A quantitative structure-activity relationship study of the skin irritation potential of phenols. <i>Toxicology in Vitro</i> 13: 915–922.</p>	<p>QSARs for skin irritation potential were studied using twenty-four phenols. Based on the hypothesis that skin irritation is induced by reaction of phenols with macromols. present in epidermal and dermal levels of the skin, the following descriptors for QSAR were selected, the abs. hardness (N) calcd. from HOMO and LUMO energy levels for reactivity, and log P (octanol-water partition coeff.) for permeability. Using these descriptors, we fitted a regression function to the set of skin irritation scores obtained from an in vivo study, which allowed derivation of equations ($r=0.85$). The equations were verified with six additional phenols, showing good correlations with the expected skin irritation scores. From the above findings, the equations can be considered useful for predicting the skin irritation potential of phenol compounds.</p>

<p>Hulzebos E, Walker JD, Gerner I & Schlegel K (2005). Use of structural alerts to develop rules for identifying chemical substances with skin irritation or skin corrosion potential. <i>QSAR Combinatorial Science</i> 24:332–342.</p>	<p>In this paper structural alerts for acute skin lesions were categorized as irritation or corrosion or a combination of corrosion/irritation alerts. Categorizing the alerts according to their mechanisms of skin irritation and corrosion and connecting them with physicochemical property limits characterizing their domain of applicability provides strategies to save test animals and costs. These alerts can be used for positive classification of chemicals causing skin irritation or skin corrosion according to EU and OECD guidelines. This paper is the third in the series of four papers describing practical, user-friendly and mechanism-based approaches for predicting when chemicals are likely to irritate or corrode the skin. In the first paper the mechanisms of skin irritation and corrosion were described. In the second paper the physicochemical property limit values for chemicals not causing skin irritation and corrosion were given. In the third paper, described here, structural alerts associated with chemicals causing skin irritation and corrosion were identified and characterized. In the fourth paper, the Skin Irritation Corrosion Rules Estimation Tool (SICRET) was described that allows users to classify chemicals as either not causing skin irritation and corrosion based on physicochemical property limit values or irritating or corrosive to the skin based on structural alerts.</p>
<p>Kodithala, K., Hopfinger, A.J., Thompson, E.D. & Robinson, M.K. (2002). Prediction of skin irritation from organic chemicals using membrane-interaction QSAR analysis. <i>Toxicological Sciences</i> 66: 336–346.</p>	<p>Membrane-interaction (MI) quantitative structure activity relationship (QSAR) analysis was carried out for a training set of 22 hydroxy organic compounds for which the Draize skin irritation scores, PII, had been determined. Significant MI-QSAR models were constructed in which skin irritation potency is predicted to increase with (1) increasing effective concentration of the compound available for uptake into phospholipid-rich regions of a cellular membrane, (2) increasing binding of the compound to the phospholipid-rich regions of a cellular membrane, and (3) the chemical reactivity of the compound as reflected by the highest occupied molecular orbital (HOMO) and/or lowest unoccupied molecular orbital (LUMO) of the molecule. Overall, the MI-QSAR models constructed for skin irritation are very similar, with respect to the types of descriptors, to those found for eye irritation. In turn, the skin irritation MI-QSAR models suggest a similar molecular mechanism of action to that postulated for eye irritation from MI-QSAR analysis. Significant and predictive QSAR models cannot be constructed unless test compound-membrane interaction descriptors are computed and used to build the QSAR models.</p>
<p>Mombelli E (2008). An evaluation of the predictive ability of the QSAR software packages, DEREK,</p>	<p>According to the REACH chemicals legislation, formally adopted by the EU in 2006, Quantitative Structure-Activity Relationships (QSARs) can be used as alternatives to</p>

<p>HAZARDEXPERT and TOPKAT, to describe chemically-induced skin irritation. ATLA 36:15–24.</p>	<p>animal testing, which itself poses specific ethical and economical concerns. A critical assessment of the performance of the QSAR models is therefore the first step toward the reliable use of such computational techniques. This article reports the performance of the skin irritation module of three commercially-available software packages: DEREK, HAZARDEXPERT and TOPKAT. Their performances were tested on the basis of data published in the literature, for 116 chemicals. The results of this study show that only TOPKAT was able to predict the irritative potential for the majority of chemicals, whereas DEREK and HAZARDEXPERT could correctly identify only a few irritant substances.</p>
<p>Rorije, E. & Hulzebos, E. (2005). Evaluation of (Q)SARs for the prediction of Skin Irritation/Corrosion Potential. Physicochemical exclusion rules. Final report for ECB contract IHCP.B430206. Joint Research Centre, Ispra, Italy. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/Evaluation_of_Eye_Irritation_QSARs.pdf</p>	<p>In this work an evaluation of a set of QSAR rules for predicting the absence of skin irritation and/or corrosion is evaluated. The results show that the evaluated rule-base is highly useful for regulatory purposes, as almost all OECD principles on (Q)SARs are met and the good (external) predictivity could lead to waiving of skin irritation tests for at least 42.3% of EU new substance notifications.</p> <p>This evaluation includes 1) the compliance of the rule-base with the OECD principles on (Q)SARs, 2) the derivation of the (Q)SAR rules, 3) the external validation of these rules, including an assessment of the suitability of the dataset used for validation. The rule-base on irritation and corrosion developed by the BfR, the German Federal Institute for Risk Assessment, predicts non-irritation and non-corrosion using physico-chemical cut-off values, defining general rules applicable to all substances and separate rules for special chemical classes of substances. The distribution of the training set data over the domains of the physico-chemical parameters used in the rule-base is visualised and analysed. Recommendations are given for setting the cut-off values of the rules at a consequent “safe” level (not allowing for any exception to the rule in the training set) and for including a consistently calculated safety margin. Specific results of the analysis were:</p> <ul style="list-style-type: none"> • The rule-base fulfils the OECD principles on (Q)SARs for the largest part • Most rules cover all irritant/corrosive substances in the training set, however: • Some physico-chemical parameters have limited predictive value: <ul style="list-style-type: none"> - lipid solubility, as this is not a generally available parameter; - vapour pressure, as the experimental data used for derivation of the rules is not conclusive, and

	<p>- melting point, as the cut-off values for this parameter were not set at a “safe” level, making predictions based on melting point less reliable.</p> <p>An external validation of the set of rules using 201 new substances not present in the training set showed 99.3% correct predictions of non-corrosivity and 96.6% correct predictions of non-irritancy. These predictions would allow declassification as R34/R35 - Corrosive for 28.4% of the chemicals, and R38 - Irritant for 42.3% of the chemicals. These results would thus allow waiving of skin irritation tests for at least 42.3% of the EU new substance notifications. Four predictions were incorrect, however, for 3 of these reasons could be given why the set of rules failed (2 substances were misclassified based on melting point rules shown to be unreliable in our analysis), and/or how these incorrect predictions can be avoided in the future. The performance of the rules then increases to 100% correct predictions of noncorrosivity and 98.8% correct predictions of non-irritancy.</p> <p>It was concluded that by applying these rules in combination with the OECD guideline 404 a large potential for saving animal lives, time and money can be realised. The rules are straightforward, easy to interpret, easily accessible and based on (measured) physicochemical data that is available for every substance that has to be notified in the EU.</p>
<p>Saliner AG, Tsakovska I, Pavan M, Patlewicz G & Worth AP (2007). Evaluation of SARs for the prediction of skin irritation/corrosion potential: structural inclusion rules in the BfR decision support system. SAR and QSAR in Environmental Research 18: 331–342.</p>	<p>The German Federal Institute for Risk Assessment (BfR) has developed a Decision Support System (DSS) to assess certain hazardous properties of pure chemicals, including skin and eye irritation/corrosion. The BfR-DSS is a rule-based system that could be used for the regulatory classification of chemicals in the European Union. The system is based on the combined use of two predictive approaches: exclusion rules based on physicochemical cut-off values to identify chemicals that do not exhibit a certain hazard (e.g., skin irritation/corrosion), and inclusion rules based on structural alerts to identify chemicals that do show a particular toxic potential. The aim of the present study was to evaluate the structural inclusion rules implemented in the BfR-DSS for the prediction of skin irritation and corrosion. The following assessments were performed: (a) a confirmation of the structural rules by rederiving them from the original training set (1358 substances), and (b) an external validation by using a test set of 200 chemicals not used in the derivation of the rules. It was found as a result that the test data set did not</p>

	match the training set relative to the inclusion of structural alerts associated with skin irritation/corrosion, albeit some skin irritants were in the test set.
Walker JD, Gerner I, Hulzebos E & Schlegel K (2004). (Q)SARs for Predicting Skin Irritation and Corrosion: Mechanisms, Transparency and Applicability of Predictions. QSAR Combinatorial Science 23: 721–725.	This paper describes previously-developed (Q)SARs for predicting skin irritation and corrosion, proposes mechanisms of skin irritation and corrosion, and discusses the transparency and applicability of predictions. This paper was written to set the tone for companion papers that describe three applications of skin irritation and corrosion (Q)SARs. The first companion paper describes physicochemical property limits that can be used to develop rules for identifying chemical substances with no skin irritation or corrosion potential. The second companion paper describes structural alerts that can be used to develop rules for identifying chemical substances with skin irritation or corrosion potential. The third companion paper describes the Skin Irritation Corrosion Rules Estimation Tool (SICRET), a user-friendly tool that allows non-(Q)SAR experts to identify chemical substances with skin irritation or corrosion potential based on physicochemical property limits and structural alerts.
Walker JD, Gerner I, Hulzebos E & Schlegel K (2005). The Skin Irritation Corrosion Rules Estimation Tool (SICRET). QSAR Combinatorial Science 24: 378–384.	This paper describes the Skin Irritation Corrosion Rules Estimation Tool (SICRET) that was developed to allow others to estimate whether their chemicals are likely to cause skin irritation or skin corrosion. SICRET uses physicochemical property limits to identify chemicals with no skin corrosion or skin irritation potential. If a chemical's physicochemical properties do not meet the prescribed limits to identify chemicals with no skin corrosion or skin irritation potential, then the chemical's structural alerts are used to identify chemicals with skin corrosion or skin irritation potential. If a chemical does not contain structural alerts that indicate it has skin corrosion or skin irritation potential, then in vitro skin corrosion or skin irritation testing is conducted. If the in vitro skin corrosion or skin irritation testing is positive, then the data are included in feedback loops for development of new structural alerts to identify chemicals with skin corrosion or skin irritation potential. If in vitro testing for skin corrosion or skin irritation is negative then the data are included in feedback loops for development of new physicochemical property limits to identify chemicals with no skin corrosion or skin irritation potential. The use of in vitro tests was proposed as a safety net to identify either new structural alerts for chemicals with skin corrosion or skin irritation potential or new physicochemical property limits for chemicals with no skin corrosion or skin irritation potential. In summary, SICRET is a tiered approach that uses physicochemical property limits, structural alerts and in-vitro tests to classify chemicals that cause skin irritation or

	skin corrosion without further animal testing.
Worth AP & Cronin MTD (2001). The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity. Toxicology 169: 119-131.	Regulatory guidelines for the assessment of acute dermal and ocular toxicity refer to the need to take the pH values of chemicals into consideration, since the acidic and basic properties of chemicals are known to play a role in the generation of acute dermal and ocular lesions. However, not all test guidelines provide an objective interpreting pH measurements in terms of acute skin or eye toxicity. The aim of this study was to develop classification models based on pH data for predicting the potential of chemicals to cause skin corrosion, skin irritation and eye irritation. The possible application of these models in the context of tiered testing strategies is discussed.

4. Serious eye damage and eye irritation

Michael Schaeffer, João Barroso, Andrew Worth & Valérie Zuang

4.1 The traditional animal test and its regulatory use

Serious eye damage is the production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application. *Eye irritation* is the production of changes in the eye following the application of a test substance to the anterior surface of the eye, which are fully reversible within 21 days of application. Currently, serious eye damage/eye irritation potential is assessed in rabbits according to OECD Test Guideline 405 (OECD, 2012), also known as the Draize eye test. The effects are graded on the basis of *in vivo* responses in the conjunctiva, cornea and iris. For classification and labelling, mean scores for each rabbit are calculated for four endpoints (corneal opacity, iritis, conjunctival chemosis, conjunctival redness) following visual grading at 24, 48 and 72 hours after installation of the test chemical to the eye sack of three different animals. A Supplement to TG 405 also includes a stepwise testing strategy for the determination of the serious eye damage/eye irritation potential. This serves to reduce and refine animal testing, since all available information on serious eye damage/eye irritation is considered prior to *in vivo* testing.

The assessment of serious eye damage/eye irritation is used for classification and labelling. The EU has implemented the United Nations Globally Harmonized System (UN, 2013) for the classification and labelling of hazardous chemicals with Directive 1272/2008 (EU CLP; EC, 2008). The test chemical is classified as “Category 1” (serious eye damage/irreversible effects on the eye) when it has the potential to seriously damage the eye, including the production of corneal, iris or conjunctival lesions in at least one animal that are not expected to reverse or have not fully reversed within an observation period of normally 21 days and/or the induction of a mean corneal opacity ≥ 3 in at least 2 of 3 animals and/or the induction of a mean iritis > 1.5 in at least 2 of 3 animals. Chemicals that have the potential to induce reversible eye irritation should be classified as “Category 2”. UN GHS provides an option (not implemented in EU CLP) to distinguish "Category 2" test chemicals into two sub-categories: “Category 2A” (irritant to eyes) when the eye irritant effects are not fully reversible within 7 days of observation; “Category 2B” (mildly irritant to eyes) when the eye irritant effects are fully reversible within 7 days of observation. Test chemicals that fall in none of the above categories do not require classification for serious eye damage/eye irritation (“No Category”). The UN scheme for the classification of serious eye damage/eye irritation based on the ocular tissue scores obtained in the Draize eye test is illustrated in Table 4.1.

Table 4.1. GHS classification of eye irritation based on observations in the Draize eye test

Location	Ocular Lesions	Grade
Cornea Opacity: degree of density	No ulceration or opacity	0
	Scattered or diffuse areas of opacity; details of iris clearly visible	1
	Easily discernible translucent area; details of iris slightly obscured	2
	Nacrous area; no details of iris visible; size of pupil barely discernible	3
	Opaque cornea; iris not discernible through the opacity	4
Iris	Normal	0
	Markedly deepened rugae, congestion, swelling, moderate circumcorneal hyperaemia; or injection; iris reactive to light (a sluggish reaction is considered to be an effect)	1
	Hemorrhage, gross destruction, or no reaction to light	2
Conjunctiva Redness: refers to palpebral and bulbar conjunctivae; excluding cornea and iris	Normal	0
	Some blood vessels hyperaemic (injected)	1
	Diffuse, crimson colour; individual vessels not easily discernible	2
	Diffuse beefy red	3
Chemosis Swelling: refers to lids and/or nictating membranes	normal	0
	Some swelling above normal	1
	Obvious swelling, with partial eversion of lids	2
	Swelling, with lids about half closed	3
	Swelling, with lids more than half closed	4
		
Decision Criteria		
<p>(a) in at least one animal effects on the cornea, iris or conjunctiva that are not expected to reverse or have not fully reversed within an observation period of normally 21 days, including grade 4 cornea lesions and other severe reactions (e.g., destruction of cornea, discoloration of the cornea by a dye substance, adhesion, pannus, interference with the function of the iris or other effects that impair sight) observed at any time during the test; and/or</p> <p>(b) in at least 2 of 3, 3 of 4, 3 of 5, or 4 of 6 tested animals, a positive response of:</p> <p>(i) corneal opacity ≥ 3, and/or</p> <p>(ii) iritis > 1.5;</p> <p>calculated as the mean scores following grading at 24, 48 and 72 hours after installation of the test chemical.</p>	<p>(a) in at least 2 of 3, 3 of 4, 3 of 5, or 4 of 6 tested animals a positive response of:</p> <p>(i) corneal opacity ≥ 1; and /or</p> <p>(ii) iritis ≥ 1; and/or</p> <p>(iii) conjunctival redness ≥ 2; and/or</p> <p>(iv) conjunctival oedema (chemosis) ≥ 2</p> <p>Calculated as the mean scores following grading at 24, 48 and 72 hours after installation of the test chemical and which fully reverses within an observation period of normally 21 days.</p>	
<p>Category 1</p> <p>Serious eye damage/Irreversible effects on the eye</p>	<p>Category 2^a</p> <p>Eye irritation/Reversible effects on the eye</p>	

^a UN GHS foresees a further categorisation of Category 2 into 2A (irritant to eyes) and 2B (mildly irritant to eyes) depending on the reversibility of the observed effects within 7 days after instillation. These optional categories are not implemented in the EU legislation.

In a recent study by Adriaens *et al.* (2014), co-sponsored by the European Commission and Cosmetics Europe, statistical resampling of *in vivo* Draize eye test data on 2134 chemicals demonstrated an overall probability of at least 11% that chemicals classified as UN GHS/EU CLP Category 1 by the Draize eye test could be equally identified as UN GHS/EU CLP Category 2 and of about 12% for UN GHS/EU CLP Category 2 chemicals to be equally identified as UN GHS/EU CLP No Category simply due to the test method's inherent within-test variability. On the other hand, the over-classification error for No Cat and Cat 2 was negligible (< 1%), which strongly suggest the high over-predictive power of the Draize eye test. Considering these results, it is probably not achievable to develop *in vitro* test methods with no false negatives. Thus, Adriaens *et al.* (2014) proposed that a false negative rate of up to 10% should be acceptable for an *in vitro* test method designed to identify chemicals not requiring classification for serious eye damage/eye irritation within UN GHS/EU CLP. Moreover, a critical revision of the UN GHS/EU CLP decision criteria for the classification of chemicals based on Draize eye test data was proposed based on the results of the analyses. In particular, it was proposed that i) the biological relevance of a threshold of 2 for classifying chemicals as Cat 2 based on conjunctival redness is reassessed considering that a significant proportion of No Category chemicals show mean conjunctival redness scores equal to or greater than 1 and that some conjunctival redness may even be present in non-treated animals; ii) conjunctival redness and/or conjunctival chemosis scores of less than 2 at day 21 are considered as fully reversed conjunctival effects and do not drive a Category 1 classification in the absence of any other Category 1 triggering effects; iii) corneal opacity scores equal to 4 that fully reverse within 21 days do not trigger a Category 1 classification in the absence of any other Category 1 triggering effects; iv) studies where a corneal opacity score equal to 4 is observed are only terminated before day 21, without investigating the reversibility of the effect and accepting a Category 1 classification, if such effects are observed in the majority of the animals tested, i.e., in 2 out of 3, 3 out of 4, 3 out of 5 or 4 out of 6 animals. Such revisions may be critical to facilitate the acceptance of alternative methods.

The results of the study by Adriaens and colleagues (Adriaens *et al.*, 2014) also indicate that the persistence and severity of corneal opacity play an equally important role in the classification of a chemical into UN GHS/EU CLP Category 1, whereas corneal opacity and conjunctival redness are the most important endpoints that determine the classification of UN GHS/EU CLP Category 2 eye irritants.

4.2 Mechanistic understanding of the endpoint

The cornea and the conjunctiva cover the anterior surface of the eye, with the cornea representing the most likely primary site of contact with chemicals. Four “mechanisms” of eye irritation were identified in an ECVAM workshop (Scott *et al.* 2010), as explained in Table 4.2.

Table 4.2. Main mechanisms of eye irritation

Mechanism / key event	Explanation
1) Cell membrane lysis	The breakdown of membrane integrity that might arise from exposure to membrane active chemicals (e.g. surfactants)
2) Coagulation	Precipitation and/or denaturation of macromolecules (proteins in particular), most likely caused by contact with acids, alkalis, or organic solvents
3) Saponification	Alkaline hydrolysis of lipids. The morphological impact may resemble those to membrane lysis and coagulation, but the effect may proceed with time through the tissue.
4) Chemical reactivity	Chemicals reacting with cellular macromolecules such as essential proteins or nucleic acids, which may lead to lysis or coagulation. This group is of particular concern because of the lack of immediate response in the tissue and thus a delayed onset of the irritation/corrosion. Examples include peroxides, mustards and bleaches.
5) Physical damage to the eye	

4.3 Status of alternative methods

Many *in vitro* test methods have been developed and evaluated for serious eye damage/eye irritation testing, as reviewed elsewhere (Eskes *et al.*, 2005; Adler *et al.*, 2010).

Since the cornea is most likely to be the primary site of contact with chemicals, corneal tissue has served as the primary model in the development of alternative methods for serious eye damage/eye irritation testing. The alternative methods can be grouped according to the nature of the test system (organotypic methods based on isolated eyes or corneas, organotypic methods based on the chicken egg chorio-allantoic membrane, reconstructed human tissues, cell-based assays, and *in chemico* assays). Some alternative test methods have gained regulatory acceptance at the EU and/or OECD levels. The main methods and their regulatory status are summarised in Table 4.3.

4.3.1 Organotypic (*ex vivo*) methods based on isolated eyes or corneas

Ex vivo methods are based on the use of isolated whole eyes or corneas. The main examples include the Bovine Corneal Opacity & Permeability test method (BCOP; Gautheron *et al.*, 1994; Gautheron, 1996), the Isolated Chicken Eye (ICE; Prinsen & Koeter, 1993; Prinsen 1994), and the Isolated Rabbit Eye (IRE; Burton *et al.*, 1981; Earl, 1994) tests. Whilst BCOP, ICE and IRE do not consider conjunctival and iridal injuries, they address corneal effects, which are the major driver of classification *in vivo* when considering the UN GHS classification. A common limitation of these methods is that they do not allow the evaluation of reversibility of corneal lesions on their own. It has been proposed, based on rabbit eye studies, that an assessment of the initial depth of corneal injury may be used to identify some types of irreversible effects (Maurer *et al.*, 2002). However, further scientific knowledge is required to understand how irreversible effects not linked with initial high level injury occur. BCOP and ICE are accepted by

the OECD to identify chemicals inducing serious eye damage (Category 1) and chemicals not requiring classification for serious eye damage/eye irritation (No Category), but they cannot be used on their own to identify Category 2 eye irritants (OECD, 2013a and b).

It is known that the under-predictions for Category 1 chemicals obtained with organotypic methods like BCOP and ICE, which were developed to detect immediate severe effects, are more likely for chemicals classified *in vivo* based only on persistence of effects than for those classified based on severity (ICCVAM, 2006; ICCVAM, 2007; ICCVAM, 2010; Adriaens *et al.*, 2014). Two organotypic methods, the Ex-Vivo Eye Irritation Test (EVEIT) based on isolated rabbit corneas (Spöler *et al.*, 2010; Frentz *et al.*, 2008) and the Porcine Cornea Reversibility Assay (PorCORA) based on isolated porcine corneas (Piehl *et al.*, 2010 and 2011), have been proposed to specifically address reversibility/persistence of effects, but neither has yet undergone a formal validation study. Both these methods are based on the direct monitoring of the recovery process in exposed excised corneas kept in culture over several days. The EVEIT uses optical coherence tomography (OCT) as a state-of-the-science, non-invasive technique to monitor the full-thickness recovery of the corneal tissue (epithelium and stroma). Methods like EVEIT and PorCORA may be needed for the correct identification of Category 1 chemicals and may significantly contribute in a testing strategy to the full replacement of the *in vivo* Draize eye test method. Histopathology may also be used in conjunction with BCOP and ICE to improve their prediction of chemicals inducing persistent effects without occurrence of initial high level injuries (classified *in vivo* based on persistence of effects) (Cazelle *et al.*, 2014). However, histopathology remains a rather subjective endpoint that only gives single time-point information as compared to methods like EVEIT and PorCORA, which are able to monitor recovery as a function of time in the living tissue. Moreover, it has been reported that although histopathology could increase the sensitivity of the organotypic assays to identify Category 1 chemicals, it may also lead to a decrease in their specificity (increase the number of chemicals over-classified as Category 1) (Schrage *et al.*, 2011).

4.3.2 Organotypic (chorio-allantoic membrane) methods

Methods based on the chicken egg chorio-allantoic membrane include the Hens Egg Test on the Chorio-Allantoic Membrane (HET-CAM; Luepke 1985; de Silva *et al.*, 1992; Gilleron *et al.*, 1996; Spielmann *et al.*, 1996) and similar methods like the Chorioallantoic Membrane Vascularization Assay (CAMVA; Bagley *et al.*, 1994). These methods have been proposed to provide information on conjunctival effects *in vivo* due to the similarity of the CAM to the conjunctiva. Both the HET-CAM and CAMVA have undergone multiple international validation studies (Bagley *et al.*, 1992b; Spielmann *et al.*, 1993; Brantom *et al.*, 1997; Spielmann *et al.*, 1996; Spielmann *et al.*, 1997; Ohno *et al.*, 1999; Hagino *et al.*, 1999; Bagley *et al.*, 1999b). More recently, HET-CAM underwent formal validation by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for the UN GHS classification system, but was not considered useful at that time to be used for regulatory purposes for the evaluation of serious eye damage/eye irritation potential of chemicals (ICCVAM, 2007; ICCVAM, 2010). ICCVAM considered that the HET-CAM might be useful for identifying chemicals not requiring hazard classification and labelling for serious eye damage/eye irritation but did not recommend the test method due to lack of data in the mild/moderate irritancy range. ICCVAM recommended that additional data be collected on mild and moderate irritants to more adequately characterise the usefulness of HET-

CAM. Following this recommendation, an international workshop on HET-CAM was held at the BfR in Berlin in October 2012 to help advance regulatory acceptance of the test method.

4.3.3 Reconstructed human Tissue (RhT) assays

Two Reconstructed human Tissue (RhT)-based test methods have recently undergone formal validation conducted by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) and Cosmetics Europe - The Personal Care Association. These are the SkinEthic™ Human Corneal Epithelium (HCE) and the EpiOcular™ Eye Irritation Test (EIT), both of which use as test systems RhTs intended to model the properties of the human corneal epithelium. Recently, the results of a prevalidation study on the EpiOcular™ EIT (Pfannenbecker *et al.* 2013) and the SkinEthic™ HCE (Alépée *et al.* 2013) were published. The Eye Irritation Validation Study (EIVS), co-sponsored by EURL ECVAM and Cosmetics Europe, evaluated the validity (relevance and reliability) of these two RhT test methods to discriminate chemicals not requiring classification for serious eye damage/eye irritancy (No Category) from chemicals requiring classification and labelling (Category 1 and Category 2) according to the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) and as implemented by the EU Classification, Labelling and Packaging regulation (EU CLP) (UN, 2013; EC, 2008). The test methods are not intended to differentiate between UN GHS/EU CLP Category 1 (serious eye damage) and UN GHS/EU CLP Category 2 (eye irritation) and were therefore not evaluated in this respect. This differentiation will need to be addressed by another tier of a test strategy (Scott *et al.*, 2010). Two protocols for each test method were evaluated in EIVS, namely the SkinEthic™ HCE Short-time Exposure (SE) and Long-time Exposure (LE) protocols and the EpiOcular™ EIT protocols for liquids and for solids. All protocols were found to be highly reproducible, but some refinements of the SkinEthic™ HCE SE and LE and of the EpiOcular™ EIT protocol for solid chemicals were deemed necessary to improve their predictive capacity. A revised EpiOcular™ EIT protocol for solid chemicals has undergone additional validation showing high reproducibility and predictive capacity. The SkinEthic™ HCE is currently undergoing external optimisation and validation outside the framework of EIVS. In the end of 2013, the European Commission submitted a project proposal to the OECD for the development of a new Test Guideline on the EpiOcular™ EIT. This project proposal was approved by the Working Group of the National Coordinators of the Test Guidelines Programme (WNT) during its April 2014 meeting and the Test Guideline is currently being drafted by the European Commission's Joint Research Centre.

4.3.4 Cell-based assays

The Fluorescein Leakage (FL) assay is the only cell-based test method that has so far gained regulatory acceptance at the OECD level (TG 460), following a retrospective validation study conducted by ECVAM. The FL can be used to identify water-soluble chemicals inducing serious eye damage (UN GHS/EU CLP Category 1). Four other cell-based assays that have been the focus of development and validation efforts are the Cytosensor Microphysiometer (CM), the Neutral Red Release (NRR), the Red Blood Cell (RBC), and the Short Time Exposure (STE) test methods. The CM, the NRR and the RBC also underwent an ECVAM-coordinated retrospective validation study. The CM was considered suitable by the Validation Management Group (VMG) and ESAC for identifying water miscible substances and mixtures inducing serious eye damage

(UN GHS/EU CLP Category 1) as well as water miscible surfactants and surfactant containing mixtures not requiring classification for serious eye damage/eye irritation (UN GHS/EU CLP No Category), and is currently under review by the OECD for adoption as a Test Guideline. The NRR test method was considered suitable for identifying water miscible chemicals not requiring classification for serious eye damage/eye irritation by the VMG, but the ESAC peer review panel (ECVAM, 2009) concluded that the available evidence was insufficient to recommend its acceptance for regulatory use. The same conclusion was reached by the VMG and ESAC on the RBC and the test method was therefore not recommended for regulatory use to predict the serious eye damage/eye irritation potential of chemicals. The STE test method underwent two validation studies, one coordinated by the Validation Committee of the Japanese Society for Alternative to Animal Experiments (JSAAE) (Sakaguchi *et al.*, 2011) and the other by the Japanese Center for the Validation of Alternative Methods (JaCVAM) (Kojima *et al.*, 2013), and was peer reviewed by NICEATM/ICCVAM. It was found suitable to identify chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) as well as surfactant solids and chemicals with a vapour pressure < 6 kPa (not highly volatile) not requiring classification for serious eye damage/eye irritation (UN GHS/EU CLP No Category). The STE test method is currently under review by the OECD for adoption as a Test Guideline.

4.3.5 *In chemico* assays

In chemico assays are based on physicochemical effects observed in acellular systems. To assess the chemical reactivity of test chemicals, the Eye Peptide Reactivity Assay (EPRA) was proposed as a variant of the Direct Peptide Reactivity Assay (DPRA), developed to screen for skin sensitizers (Gerberick *et al.*, 2004). It is based on the reactivity of a test chemical with Cysteine and Lysine containing heptapeptides. The EPRA has been used in a tiered test strategy together with the SkinEthic™ HCE test method as a means of choosing between two treatment procedures in the SkinEthic™ HCE test method (short exposure for reactive chemicals and long exposure for non-reactive chemicals; Alépée *et al.*, 2013).

Another example is the Ocular Irritation® assay which predicts the ocular irritation potential of chemicals by measuring protein denaturation. This test method mimics the biochemical phenomena of corneal protein denaturation and disruption caused by irritant chemicals acting on the cornea. The chemical effects observed in the Ocular Irritation® assay are assumed to mimic the disruptive effects ocular irritants can have on the highly organised structure of corneal proteins and carbohydrates which result in corneal cloudiness/opacity in the *in vivo* Draize eye test. This assay underwent an external validation study and is being proposed to identify chemicals not requiring classification for serious eye damage/eye irritation (UN GHS No Category) and chemicals inducing serious eye damage (UN GHS Category 1) (Eskes *et al.*, 2014).

4.3.6 *Quantitative Structure-Activity Relationships*

Various QSAR modelling studies have attempted to model eye irritation responses, focusing either on the prediction of weighted averages of the underlying tissue scores, such as the MMAS (modified maximum average score; e.g. Abraham *et al.*, 2003), or on classification models (e.g. Worth & Cronin, 1999). The status of literature-based models has changed little since the previous JRC reviews (Gallegos Saliner *et al.*; 2006,

2008), with relatively few new publications appearing since 2008 (e.g. Solimeo *et al.*, 2012).

Software tools with functionalities for predicting eye irritation potential include Toxtree, the OECD QSAR Toolbox, Derek Nexus, TOPKAT, Molcode QSARModel, and Multi-CASE. The freely available Toxtree and the OECD Toolbox tools include the BfR rulebase which applies a set of physicochemical rules to predict the absence of effects, along with a set of structural rules to predicting the presence of effects. Evaluations by the JRC of the predictive performance of the physicochemical (Tsakovska *et al.*, 2005) and structural (Tsakovska *et al.*, 2007) rules found a high negative predictivity (87%) of the physicochemical exclusion rules, and a high positive predictivity of the structural inclusion rules (>80%, depending on the hazard class). These evaluations were performed before the BfR rulebase was implemented in Toxtree and the OECD Toolbox, so it would be worthwhile to revisit the analysis, using the modern tools and more recently published data.

These software tools can be used in a WoE approach or tiered testing strategy. As with all computational tools, the predictions need to be carefully evaluated using information on the model characteristics. For classification and labelling, the BfR rulebase provides information that is closest to the regulatory goal, since the system was designed to predict the then EU Risk Phrases for irritation (R36, R41) and corrosion (R34, R35).

4.4 Status of alternative testing strategies

In the area of toxicological assessment for eye hazards, it is currently generally accepted that, in the foreseeable future, no single *in vitro* serious eye damage/eye irritation test method will be able to replace the *in vivo* Draize eye test to predict across the full range of ocular effects for different chemical classes. However, strategic combinations of several alternative methods (*in vitro* and *in silico*) may be able to replace the Draize eye test. Testing strategies for serious eye damage/eye irritation have been reviewed (Gallegos Saliner & Worth, 2007) or proposed (Grindon *et al.*, 2008) elsewhere. More recently, a conceptual framework for tiered testing strategies was developed within an ECVAM workshop (Scott *et al.*, 2010). The framework is based on alternative serious eye damage/eye irritation methods that vary in their capacity to detect either chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) or chemicals not requiring classification for serious eye damage/eye irritation (UN GHS/EU CLP No Category). According to this framework, the entire range of irritancy may be resolved by arranging tests in a tiered strategy that may be operated from either end. If the test chemical is expected to be a moderate to severe eye irritant, the “Top-Down approach” is initiated, in which chemicals inducing serious eye damage are identified first. Conversely, if the test chemical is expected to be a non-irritant or mild irritant, the “Bottom-Up” approach is initiated, starting with the identification of non-irritants (Figure 4.1). The Top-Down/Bottom-Up framework thus proposes a two-step procedure where validated methods are used to determine if a test chemical is not classified (UN GHS/EU CLP No Category) or if it induces serious eye damage (UN GHS/EU CLP category 1). If neither a No Category nor a Category 1 is assigned in this stepwise tiered testing, the framework proposes that either an *in vivo* test is conducted or a default UN GHS/EU CLP Category 2 classification is assigned (Figure 4.1).

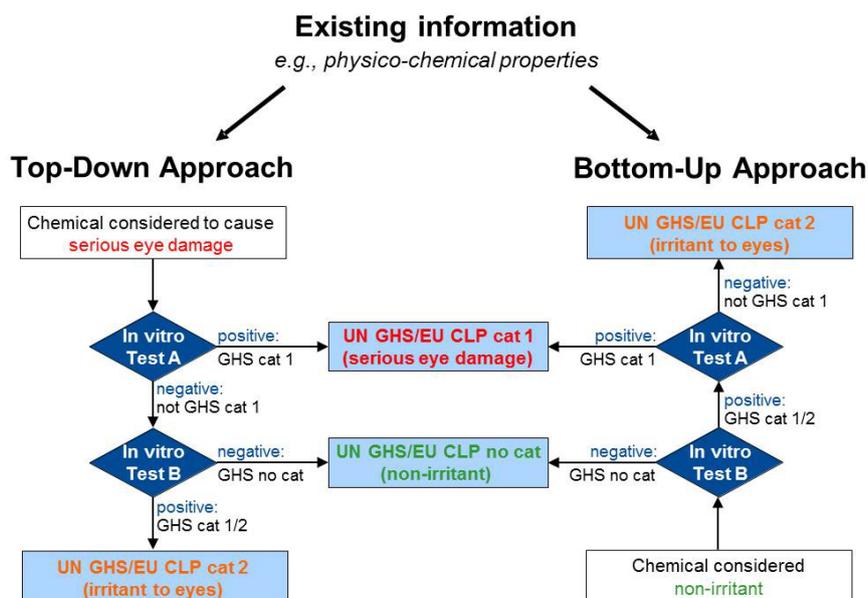


Figure 4.1. Bottom-Up and Top-Down approaches for serious eye damage/eye irritation (after Scott *et al.*, 2010)

One of the problems associated with the two-tier Top-Down/Bottom-Up testing strategy is that a default UN GHS/EU CLP Category 2 classification after only testing in two test methods will generate a significant number of false negative (Category 1 underclassified as Category 2) and false positive (No Category overclassified as Category 2) results. On the one hand, currently accepted methods for identifying UN GHS/EU CLP Category 1, like BCOP and ICE, underpredict 14-48% of the *in vivo* Category 1 chemicals, mostly those inducing persistent effects without occurrence of initial high level injuries (classified *in vivo* based only on persistence of effects). This shortcoming may be resolved by inclusion in the strategy of a method specifically addressing reversibility/persistence of effects based on the direct monitoring of the recovery process, like EVEIT or PorCORA (Figure 4.2). On the other hand, it is clear that due to the very high sensitivity required by regulatory authorities for accepting the use of *in vitro* test methods to identify chemicals not requiring hazard classification and labelling for serious eye damage/eye irritation (UN GHS/EU CLP No Category), their specificity will never go beyond 60-80% (the highest the specificity, the more limited the applicability). EpiOcular™ EIT, ICE and STE are the three test methods showing the best accuracy for identifying UN GHS/EU CLP No Category chemicals and their specificity is only 60-80% with already a few false negatives being obtained (sensitivity around 95%). In such a scenario, several methods capable of identifying UN GHS/EU CLP No Category chemicals with very high sensitivity will need to be combined to increase the overall specificity of the testing strategy to acceptable values (Figure 4.3).

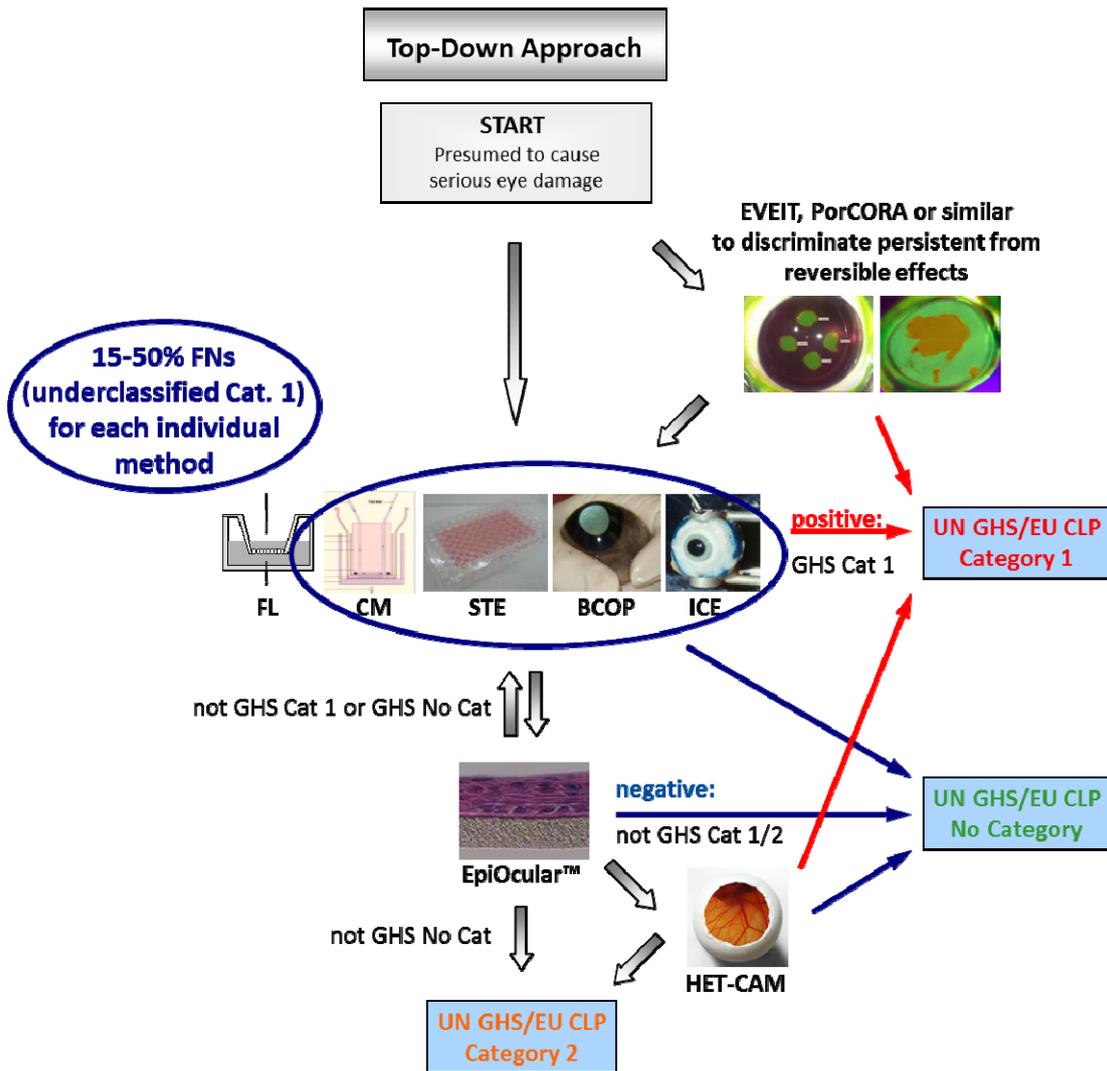


Figure 4.2. Illustrated example of a Top-Down testing strategy for serious eye damage/eye irritation combining methods like EpiOcular™ EIT, BCOP, ICE, FL, STE, CM, HET-CAM, EVEIT and/or PorCORA. The order of the different methods in the strategy presented is only hypothetical and it will depend on several aspects like the purpose of the testing, the complementarity of the different test methods, their limitations, their price, their rates of false positives and false negatives, and their mechanistic coverage. EpiOcular™ EIT, STE, CM, HET-CAM, EVEIT and PorCORA are still not covered by an OECD Test Guideline and therefore still need to gain regulatory acceptance. ICE image courtesy of Dr Menk Prinsen, TNO, The Netherlands; BCOP image courtesy of Institute for In Vitro Sciences, Inc., USA; EpiOcular™ image courtesy of MatTek Corporation, USA; EVEIT image courtesy of Dr Norbert Schrage, ACTO e.V. & Dr. Felix Spöler, IHT, RWTH Aachen University, Germany; PorCORA image courtesy of MB Research Laboratories, USA)

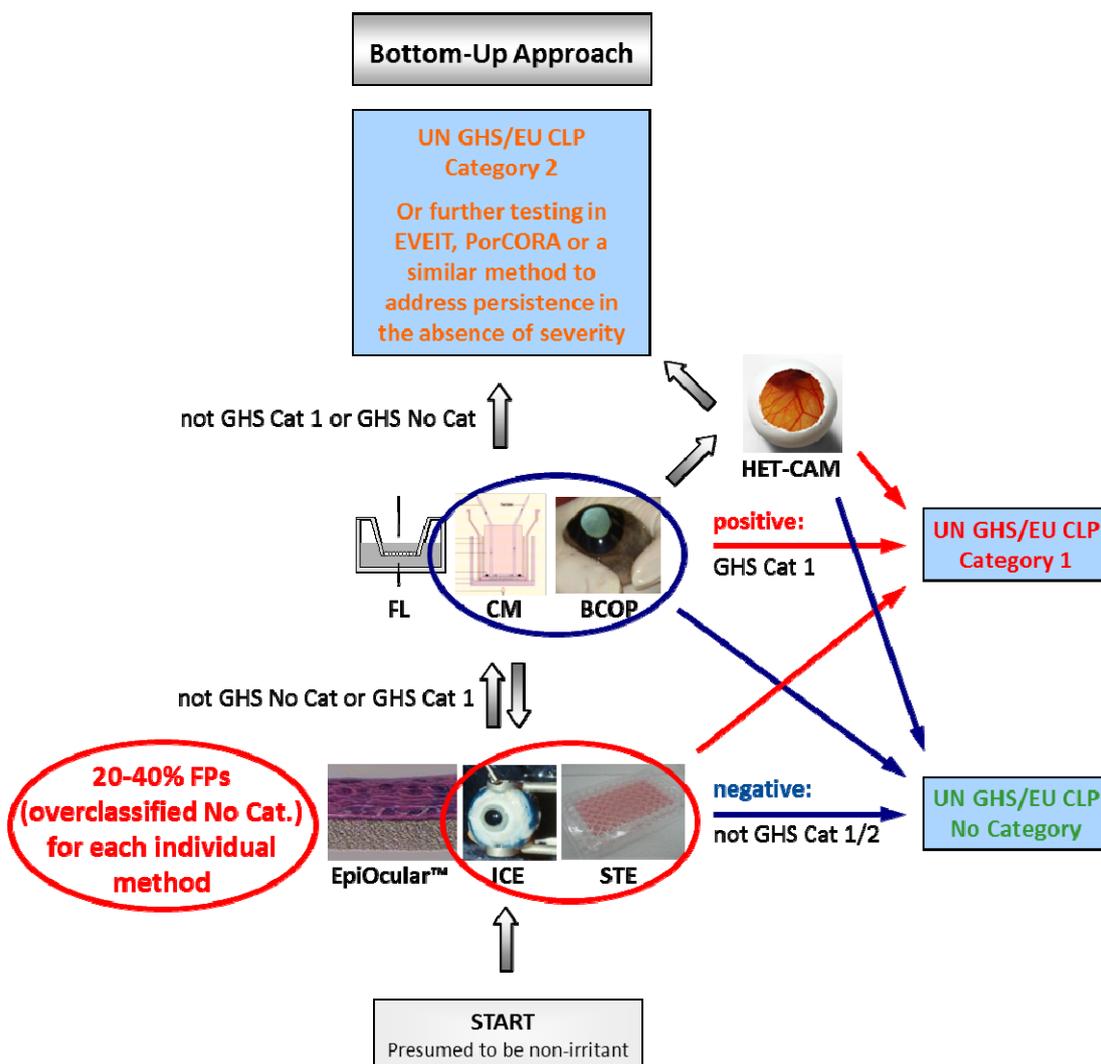


Figure 4.3. Illustrated example of a Bottom-Up testing strategy for serious eye damage/eye irritation combining methods like EpiOcular™ EIT, BCOP, ICE, FL, STE, CM, HET-CAM, EVEIT and/or PorCORA. The order of the different methods in the strategy presented is only hypothetical and it will depend on several aspects like the purpose of the testing, the complementarity of the different test methods, their limitations, their price, their rates of false positives and false negatives, and their mechanistic coverage. EpiOcular™ EIT, STE, CM, HET-CAM, EVEIT and PorCORA are still not covered by an OECD Test Guideline and therefore still need to gain regulatory acceptance. ICE image courtesy of Dr Menk Prinsen, TNO, The Netherlands; BCOP image courtesy of Institute for In Vitro Sciences, Inc., USA; EpiOcular™ image courtesy of MatTek Corporation, USA

Potential ways of combining *in vitro* tests methods in testing strategies based on the concept of the Bottom-up and Top-down approaches have been investigated by Kolle *et al.* (2011) and Hayashi *et al.* (2012). Kolle *et al.* (2011) combined EpiOcular™ EIT and BCOP in a two-tier Bottom-up/Top-Down test strategy and Hayashi *et al.* (2012) combined EpiOcular™ EIT, BCOP, STE and HET-CAM in a two-stage Bottom-Up tiered approach. From the data presented in these two publications, it can be seen that

the specificity for identifying UN GHS/EU CLP No Category chemicals will increase substantially (to close to 90%) by combining in a test strategy several methods able to identify this type of chemicals. This occurs as a result of multiple methods complementing each other by correctly identifying different sets of UN GHS/EU CLP No Category chemicals. Importantly, the increase in specificity of the test strategy as compared to the individual methods is not accompanied by a significant decrease in sensitivity due to the very high sensitivity already displayed by all of these methods on their own (> 90%). In the data published by Kolle *et al.* (2011), it can be seen that one UN GHS/EU CLP No Category chemical is overpredicted by EpiOcular™ EIT but correctly predicted by BCOP, while 10 other UN GHS No Category chemicals are overpredicted by BCOP but correctly predicted by EpiOcular™ EIT. Hayashi *et al.* published results obtained with BCOP, EpiOcular™ EIT, STE and HET-CAM on 26 UN GHS/EU CLP No Category chemicals, 13 UN GHS/EU CLP Category 2 chemicals and 16 UN GHS/EU CLP Category 1 chemicals. Ten of the 26 UN GHS/EU CLP No Category chemicals were correctly identified by BCOP, 13 by EpiOcular™ EIT, 17 by STE, and 9 by HET-CAM, which gives individual false positive rates ranging from 35% to 65%. However, for 23 of these UN GHS/EU CLP No Category chemicals (88.5%) a correct prediction was obtained in at least one of these 4 methods. If one would start a Bottom-Up approach with EpiOcular™ EIT, 13 of the UN GHS/EU CLP No Category chemicals would be overpredicted as irritant (50% false positives) and no false negatives would be produced. Using BCOP in the next tier as a method able to identify both Category 1 and No Category, the latter with 100% sensitivity (no false negatives), two of the EpiOcular™ EIT false positives would be correctly identified by BCOP as UN GHS/EU CLP No Category without producing any false negatives. Fifteen of the 16 UN GHS/EU CLP Category 1 chemicals would also be correctly identified by BCOP, with the one Category 1 that is missed being a chemical that is classified *in vivo* based only on persistence of effects. Finally, 7 of the 13 UN GHS/EU CLP Category 2 chemicals would be overclassified as Category 1. If HET-CAM would then be used in a third tier to further identify UN GHS/EU CLP No Category chemicals it would correctly identified two other EpiOcular™ EIT false positives as No Category. HET-CAM alone would underclassify 1 of the 29 classified chemicals (a Category 2A) as No Category, but in a test strategy, this chemical would have been already overclassified as Category 1 by BCOP. Finally, using STE as the final tier, a further 6 of the EpiOcular™ EIT false positives would be correctly identified as No Category, but in this case two false negatives would be produced, one Category 2A and one Category 2B. STE alone would actually underclassify 4 of the 29 classified chemicals (3 Category 2A and 1 Category 2B) as No Category, but in a test strategy, two of these chemicals (2 Category 2A) would have been already overclassified as Category 1 by BCOP. In conclusion, by combining multiple methods able to identify UN GHS No Category chemicals, the number of correctly identified UN GHS/EU CLP No Category chemicals would increase from 35-65% to 88.5%, and at the same time the accuracy for the identification of UN GHS/EU CLP Category 2 by default at the end of the strategy would be significantly improved.

4.5 Conclusions

Information on serious eye damage/eye irritation potential, traditionally obtained by the Draize rabbit eye test, is required for the purposes of REACH, the Biocidal Products Regulation and the CLP regulation. While it is generally accepted that no single alternative method for serious eye damage/eye irritation will be able to replace the *in*

vivo Draize eye test across the full range of irritation for different chemical classes, strategic combinations of several alternative test methods may be able to replace the Draize eye test. With more and more alternative tools becoming available, Integrated Approaches to Testing and Assessment (IATA), incorporating existing information, QSAR predictions, and new *in vitro* test results, should now be developed to provide a means for standardised and comprehensive decision making.

Testing strategies such as the Top-Down/Bottom-Up approach provide a convenient framework for generating new *in vitro* test data when a decision on classification and labelling cannot be taken based on existing data and QSAR predictions. The framework is based on the combination of *in vitro* methods able to identify chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) and/or chemicals not requiring classification for serious eye damage/eye irritation (UN GHS/EU CLP No Category). A number of OECD Test Guidelines based on *in vitro* test methods have already been adopted (TG 437, 438, 460), and three other draft Test Guidelines on CM, STE and EpiOcular™ EIT are under discussion at OECD level. All these Test Guidelines identify chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) and/or chemicals not classified for serious eye damage/eye irritation in the framework of a Top-Down/Bottom-Up approach. In the absence of suitable methods for directly identifying eye irritants (UN GHS/EU CLP Category 2), the Top-Down/Bottom-Up approach may be used to indirectly identify such chemicals by an accurate exclusion of Category 1 and No Category. For this framework to work with acceptable accuracy, several methods addressing each of the two ends of the irritancy scale will however need to be combined. Although combinations of the existing assays within testing strategies such as described above are already encouraging, a full replacement of the *in vivo* Draize eye test has still not been achieved, mostly because of the lack of test methods able to identify Category 2 chemicals and to discriminate persistent from reversible effects.

Further research efforts will need to focus on the development of mechanistically-based *in vitro* assays and, in particular, methods capable of detecting conjunctival effects (important in the classification of Category 2 eye irritants) as well as the persistence of effects (important in the classification of Category 1 chemicals) (Adriaens *et al.*, 2014). In addition, it would be worthwhile to re-evaluate some of the QSAR tools with a view to applying them in an IATA for serious eye damage/eye irritation.

4.6 References

- Abraham MH, Hassanisadi M, Jalali-Heravi M, Ghafourian T, Cain WS & Cometto-Muniz JE (2003). Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis. *Toxicological Sciences* 76: 384-391.
- Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleijnans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485.
- Adriaens E (2010). Development of a testing strategy for eye irritation testing as alternative to the *in vivo* Draize test. Final Report to JRC-ECVAM. JRC Contract No. CCR.IHCP.C433251.XO.
- Adriaens E, Barroso J, Eskes C, Hoffmann S, McNamee P, Alépée N, Bessou-Touya S, De Smedt A, De Wever B, Pfannenbecker U, Tailhardat M & Zuang V (2014). Retrospective analysis of the Draize test for serious eye damage/eye irritation: importance of understanding the *in vivo* endpoints under UN GHS/EU CLP for the development and evaluation of *in vitro* test methods. *Archives of Toxicology* 88: 701-723.
- Alépée N, Bessou-Touya S, Cotovio J, De Smedt A, De Wever B, Faller C, Jones P, Le Varlet B, Marrec-Fairley M, Pfannenbecker U, Tailhardat M, Van Goethem F & McNamee P (2013). Cosmetics Europe Multi-Laboratory Pre-Validation of the SkinEthic™ Reconstituted Human Corneal Epithelium Test Method for the Prediction of Eye Irritation. *Toxicology in Vitro* 27: 1476-1488.
- Bagley DM, Bruner LH, de Silva O, Cottin M, O'Brien KAF, Uttley M & Walker AP (1992). An evaluation of five potential alternatives *in vitro* to the rabbit eye irritation test *in vivo*. *Toxicology in Vitro* 6: 275-284.
- Bagley DM, Waters D & Kong BM (1994). Development of a 10-day chorioallantoic membrane vascular assay as an alternative to the Draize rabbit eye irritation test. *Food Chem. Toxicol.* 32: 1155-1160.
- Bagley DM, Cerven D & Harbell J (1999). Assessment of the chorioallantoic membrane vascular assay (CAMVA) in the COLIPA *in vitro* eye irritation validation study. *Toxicology in Vitro* 13: 285-293.
- Brantom PG, Bruner LH, Chamberlain M, de Silva O, Dupuis J, Earl LK, Lovell DP, Pape WJ, Uttley M, Bagley DM, Baker FW, Bracher M, Courtellemont P, Declercq L, Freeman S, Steiling W, Walker AP, Carr GJ, Dami N, Thomas G, Harbell J, Jones PA, Pfannenbecker U, Southee JA, Tchong M, Argembeaux H, Castelli D, Clothier R, Esdaile DJ, Itigaki H, Jung K, Kasai Y, Kojima H, Kristen U, Larnicol M, Lewis RW, Marenus K, Moreno O, Peterson A, Rasmussen ES, Robles C & Stern M (1997). A summary report of the COLIPA international validation study on alternatives to the draize rabbit eye irritation test. *Toxicology in Vitro* 11: 141-179.

Burton ABG, York M & Lawrence RS (1981). The in vitro assessment of severe eye irritants. *Food and Cosmetics Toxicology* 19: 471-480.

Cazelle E, Eskes C, Hermann M, Jones P, McNamee P, Prinsen M, Taylor H & Wijnands MV (2014). Suitability of histopathology as an additional endpoint to the Isolated Chicken Eye Test for classification of non-extreme pH detergent and cleaning products. *Toxicology in Vitro* 28: 657-666.

Chamberlain M, Gad SC, Gautheron P & Prinsen MK (1997). IRAG (Interagency Regulatory Alternatives Group) working group 1. Organotypic models for the assessment/prediction of ocular irritation. *Food and Chemical Toxicology* 35: 23-37.

De Silva O, Rougier A & Dossou KG (1992). The HET-CAM test: a study of the irritation potential of chemicals and formulations. *ATLA* 20: 432-437.

Draize JH, Woodard G & Calvery HO (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *Journal of Pharmacology and Experimental Therapeutics* 82: 377-390.

Earl L (1994). The rabbit enucleated eye test. *Invitro* Protocol 85.

EC (2004). Manual of Decisions for Implementation of the 6th and 7th Amendments to Directive 67/548/EEC on Dangerous Substances. Updated version of July 2004 (EUR 20519). Ispra: European Chemicals Bureau, European Commission JRC, 2004.

EC (2008). Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (2008). Available from: <http://ec.europa.eu/enterprise/sectors/chemicals/documents/classification/#h2-1>.

ECVAM (2007). ESAC statement on the conclusion of the ICCVAM retrospective study on Organotypic in vitro assays as screening tests to identify potential ocular corrosives and severe irritants as determined by US EPA, EU(R41) and UN GHS classifications in a tiered testing strategy, as part of a weight of evidence approach. 27 April 2007. ECVAM Scientific Advisory Committee. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/publication/ESAC26_statement_Organotypic_20070510_C.pdf.

ECVAM (2009). Statement on the Scientific Validity of Cytotoxicity/Cell Function Based In Vitro Assays For Eye Irritation Testing. 10 July 2009. ECVAM Scientific Advisory Committee. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/validation-regulatory-acceptance/docs-eye-irritation/ESAC31_CBA_eye-irritation_20091005.pdf.

Eskes C, Bessou S, Bruner L, Curren R, Harbell J, Jones P, Kreiling R, Liebsch M, McNamee P, Pape W, Prinsen MK, Seidle T, Vanparrys P, Worth A & Zuang V. (2005). Eye irritation. *ATLA Alternatives to Laboratory Animals*, 33 (SUPPL. 1), pp. 47-81.

Eskes C, Hoffmann S, Facchini D, Ulmer R, Wang A, Flego M, Vassallo M, Bufo M, van Vliet E, d'Abrosca F & Wilt N (2014). Validation study on the Ocular Irritation® assay for eye irritation testing. *Toxicology in Vitro* 28: 1046-1065.

Freeman SJ, Alépée N, Barroso J, et al. (2010) Prospective validation study of reconstructed human tissue models for eye irritation testing. *ALTEX* 27, Special Issue 2010: 261-266.

Frentz M, Goss M, Reim M, & Schrage NF (2008). Repeated exposure to Benzalkonium chloride in the Ex Vivo Eye Irritation Test (EVEIT): observation of isolated corneal damage and healing. *ATLA* 36: 25-32.

Gautheron P, Giroux J, Cottin M, Audegond L, Morilla A, Mayordomo-Blanco L, Tortajada A, Haynes G & Vericat JA (1994). Interlaboratory assessment of the bovine corneal opacity and permeability (BCOP) assay. *Toxicology in Vitro* 8: 381-392.

Gautheron P (1996). The bovine corneal opacity and permeability assay. *Invitox Protocol* 98.

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW & Lepoittevin JP (2004). Development of a peptide reactivity assay for screening contact allergens. *Toxicological Sciences* 81: 332-343.

Gilleron L, Coecke S, Sysmans M, Hansen E, Van Oproy S, Marzin D, Van Cauteren H & Vanparys P (1996). Evaluation of a modified HET-CAM assay as a screening test for eye irritancy. *Toxicology in Vitro* 10: 431-446.

Grindon C, Combes R, Cronin MT, Roberts DW & Garrod JF (2008). An integrated decision-tree testing strategy for eye irritation with respect to the requirements of the EU REACH legislation. *ATLA* 6: 81-92.

Hagino S, Kinoshita S, Tani N, Nakamura T, Ono N, Konishi K, Iimura H, Kojima H & Ohno Y (1999). Interlaboratory validation of in vitro eye irritation tests for cosmetic ingredients. (2) Chorioallantoic membrane (CAM) test. *Toxicology in Vitro* 13: 99-113.

Hayashi K, Mori T, Abo T, Ooshima K, Hayashi T, Komano T, Takahashi Y, Sakaguchi H, Takatsu A & Nishiyama N (2012). Two-stage bottom-up tiered approach combining several alternatives for identification of eye irritation potential of chemicals including insoluble or volatile substances. *Toxicol In Vitro* 26: 1199-1208.

ICCVAM (2006). Background Review Documents: Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA. Available from: http://iccvam.niehs.nih.gov/methods/ocudocs/ocu_brd.htm.

ICCVAM (2007). ICCVAM Test Method Evaluation Report: *In Vitro* Ocular Toxicity Test Methods for Identifying Severe Irritants and Corrosives. NIH Publication No. 07-4517. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA. Available from: http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm.

ICCVAM (2010). ICCVAM Test Method Evaluation Report: Current Validation Status of In Vitro Test Methods Proposed for Identifying Eye Injury Hazard Potential of Chemicals and Products. NIH Publication No. 10-7553. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA. Available from: <http://iccvam.niehs.nih.gov/methods/ocutox/MildMod-TMER.htm>.

JRC (2004). Manual of Decisions for Implementation of the 6th and 7th Amendments to Directive 67/548/EEC on Dangerous Substances. Updated version of July 2004 (EUR 20519). Ispra: European Chemicals Bureau, European Commission Joint Research Centre.

Kojima H, Hayashi K, Sakaguchi H, Omori T, Otoizumi T, Sozu T, Kuwahara H, Hayashi T, Sakaguchi M, Toyoda A, Goto H, Watanabe S, Ahiko K, Nakamura T & Morimoto T (2013).

Second-phase validation study of Short Time Exposure test for assessment of eye Irritation potency of chemicals. *Toxicol. In Vitro* 27: 1855-1869.

Luepke NP (1985). Hen's egg chorioallantoic membrane test for irritation potential. *Food and Chemical Toxicology* 23: 287-291.

Maurer JK, Parker RD & Jester JV (2002). Extent of corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. *Reg. Tox. Pharmacol.* 36: 106-117.

Nguyen DH, Beuerman RW, De Wever B & Rosdy M (2003). Three-dimensional construct of the human corneal epithelium for in vitro toxicology. In *Alternatives Toxicological Methods*. Salem, H. and Katz S.A., Eds. CRC press, 47-159.

OECD (2012a). Test Guideline 460. Fluorescein Leakage Test Method for Identifying Ocular Corrosives and Severe Irritants, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264185401-en. Available from: <http://www.oecd.org/env/testguidelines>.

OECD (2012b). Test Guideline 405. OECD Guideline for the Testing of Chemicals: Acute Eye Irritation/Corrosion. Paris, France: Organisation for Economic Cooperation and Development. Section 4, OECD Publishing. doi: 10.1787/9789264070646-en. Available from: <http://www.oecd.org/env/testguidelines>.

OECD (2013a). Test Guideline 437. Bovine Corneal Opacity and Permeability Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076303-en. Available from: <http://www.oecd.org/env/testguidelines>.

OECD (2013b). Test Guideline 438. Isolated Chicken Eye Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076310-en. Available from: <http://www.oecd.org/env/testguidelines>.

Ohno Y, Kaneko T, Inoue T, Morikawa Y, Yoshida T, Fujii A, Masuda M, Ohno T, Hayashi M, Momma J, Uchiyama T, Chiba K, Ikeda N, Imanishi Y, Itakagaki H, Kakishima H, Kasai Y, Kurishita A, Kojima H, Matsukawa K, Nakamura T, Ohkoshi K, Okumura H, Saijo K, Sakamoto K, Suzuki T, Takano K, Tatsumi H, Tani N, Usami M & Watanabe R (1999). Interlaboratory validation of the in vitro eye irritation tests for cosmetic ingredients. (1) Overview of the validation study and Draize scores for the evaluation of the tests. *Toxicology in Vitro* 13: 73-98.

Piehl M, Gilotti A, Donovan A, DeGeorge G & Cerven D (2010). Novel cultured porcine corneal irritancy assay with reversibility endpoint. *Toxicology in Vitro* 24: 231-239.

Piehl M, Carathers M, Soda R, Cerven D & DeGeorge G (2011). Porcine Corneal Ocular Reversibility Assay (PorCORA) predicts ocular damage and recovery for global regulatory agency hazard categories. *Toxicology in Vitro* 25: 1912-1918.

Pfannenbecker U, Bessou-Touya S, Faller C, Harbell J, Jacob T, Raabe H, Tailhardat M, Alépée N, De Smedt A, De Wever B, Jones P, Kaluzhny Y, Le Varlet B, McNamee P, Marrec-Fairley M & Van Goethem F (2013). Cosmetics Europe multi-laboratory pre-validation of the

EpiOcular™ reconstituted human tissue test method for the prediction of eye irritation. *Toxicology in Vitro* 27: 619-626.

Prinsen MK & Koeter HBWM (1993). Justification of the Enucleated Eye Test with eyes of slaughterhouse animals as an alternative to the Draize Eye Irritation Test with rabbits. *Food and Chemical Toxicology* 31: 69-76.

Prinsen M (1994). Chicken enucleated eye test (CEET). Invitox Protocol 80.

Sakaguchi H, Ota N, Omori T, Kuwahara H, Sozu T, Takagi Y, Takahashi Y, Tanigawa K, Nakanishi M, Nakamura T, Morimoto T, Wakuri S, Okamoto Y, Sakaguchi M, Hayashi T, Hanji T & Watanabe S (2011). Validation study of the Short Time Exposure (STE) test to assess the eye irritation potential of chemicals, *Toxicology in Vitro* 25: 796-809.

Schrage A, Kolle SN, Moreno MC, Norman K, Raabe H, Curren R, van Ravenzwaay B & Landsiedel R (2011). The bovine corneal opacity and permeability test in routine ocular irritation testing and its improvement within the limits of OECD test guideline 437. *Altern Lab Anim.* 39: 37-53.

Scott L, Eskes C, Hoffmann S, et al. (2010) A proposed eye irritation testing strategy to reduce and replace in vivo studies using Bottom-Up and Top-Down approaches. *Toxicology in Vitro* 24: 1-9.

Solimeo R, Zhang J, Kim M, Sedykh A & Zhu H (2012). Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach. *Chemical Research in Toxicology* 25: 2763-2769

Spielmann H, Kalweit S, Liebsch M, Wirnsberger T, Gerner I, Bertram-Neis E, Krauser K, Kreiling R, Miltenburger HG, Pape W & Steiling W (1993). Validation study of alternatives to the Draize eye irritation test in Germany: Cytotoxicity testing and HET-CAM test with 136 industrial chemicals. *Toxicology in Vitro* 7: 505-510.

Spielmann H, Liebsch M, Kalweit S, Moldenhauer F, Wirnsberger T, Holzhütter HG, Schneider B, Glaser S, Gerner I, Pape WJW, Kreiling R, Krauser K, Miltenburger HG, Steiling W, Luepke NP, Müller N, Kreuzer H, Mürmann P, Spengler J, Bertram-Neis E, Siegemund B & Wiebel FJ (1996). Results of a validation study in Germany on two in vitro alternatives to the Draize eye irritation test, HET-CAM test and the 3T3 NRU cytotoxicity test. *ATLA* 24: 741–858.

Spielmann H, Liebsch M, Moldenhauer F, Holzhütter HG, Bagley DM, Lipman JM, Pape WJ, Miltenburger H, de Silva O, Hofer H & Steiling W (1997). IRAG working group 2. CAM-based assays. Interagency Regulatory Alternatives Group. *Food and Chemical Toxicology* 35: 39-66.

Spöler F, Frenz M & Schrage NF (2010). Towards a new in vitro model of dry eye: the Ex Vivo Eye Irritation Test. *Dev Ophthalmol.* 45: 93-107.

Tsakovska I, Netzeva T & Worth AP (2005). Evaluation of (Q)SARs for the prediction of Eye Irritation/Corrosion Potential - physicochemical exclusion rules. European Commission Report EUR 21897 EN. European Commission-Joint Research Centre, Ispra, Italy. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology

Tsakovska I, Gallegos A, Netzeva T, Pavan M & Worth AP (2007). Evaluation of SARs for the prediction of eye irritation/corrosion potential – structural inclusion rules in the BfR decision support system. *SAR & QSAR in Environmental Research* 18: 221-235.

UN (2013) Globally Harmonized System of Classification and Labelling of Chemicals (GHS), 5th revised edition. United Nations, New York and Geneva, 2013. Available from:

http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf.

Worth AP & Cronin MTD (1999). Embedded Cluster Modelling – a novel method for analysing embedded data sets. *Quantitative Structure-Activity Relationships* 18: 229-235.

Table 4.3. Alternative methods for serious eye damage/eye irritation

Method	Test system	Description	Applicability	Regulatory status
<i>Organotypic (Ex vivo isolated eyes) methods</i>				
Bovine Corneal Opacity & Permeability (BCOP)	<i>Ex vivo</i> bovine cornea, obtained as a by-product from abattoirs	Based on normal physiological and biochemical function of the bovine cornea <i>in vitro</i> . The damage caused by the test chemical is assessed by quantitative measurements of changes in corneal opacity and permeability with an opacitometer and a visible light spectrophotometer. Permeability is quantified as the amount of fluorescein dye that passes across the full thickness of the cornea from the epithelial surface.	Suitable for identifying UN GHS/EU CLP Category 1 (JRC, 2004; ECVAM, 2007) and UN GHS/EU CLP No Category chemicals. The false positive rate for alcohols and ketones and the false negative rates for solids have been noted as limitations for the identification of UN GHS/EU CLP Category 1 (OECD, 2013a).	OECD TG 437 (adopted in 09/2009 and updated in 07/2013).
Isolated Chicken Eye test (ICE)	<i>Ex vivo</i> whole chicken eye, obtained as a by-product from abattoirs	Damage by the test substance is assessed by determination of corneal swelling, opacity, and fluorescein retention, which is a subjective measurement of the degree of the fluorescein dye that is retained by epithelial cells in the cornea after the exposure to a test chemical. The extent of fluorescein retention is representative of damage to the corneal epithelium. While the latter two parameters involve a qualitative assessment, analysis of corneal swelling provides for a quantitative assessment. Each measurement is either converted into a quantitative score used to calculate an overall Irritation Index, or assigned a qualitative categorization that is used to	Suitable for identifying UN GHS/EU CLP Category 1 (JRC, 2004; ECVAM, 2007) and UN GHS/EU CLP No Category chemicals. The false positive rate for alcohols and the false negative rates for solids and surfactants have been noted as limitations for the identification of UN	OECD TG 438 (adopted in 09/2009 and updated in 07/2013).

Method	Test system	Description	Applicability	Regulatory status
		assign an <i>in vitro</i> classification. Either of these outcomes can then be used to predict the <i>in vivo</i> ocular hazard effects of a test chemical.	GHS/EU CLP Category 1 (OECD, 2013b).	
Isolated Rabbit Eye test (IRE)	<i>Ex vivo</i> whole rabbit eye	<p>Determines the opacification of the cornea and the increase in corneal thickness (corneal swelling) after exposure to irritant substances. Whole eyeballs obtained by immediate dissection from humanely killed laboratory rabbits with healthy eyes are mounted and maintained in a vertical position in a so-called superfusion chamber with controlled temperature and humidity.</p> <p>Primarily predicts corneal effects and does not provide information on the effects of chemicals on the conjunctiva of the eye or on the recovery of the cornea from damage (beyond a few hours), which might occur in the eye <i>in vivo</i>.</p>	Suitable for identifying UN GHS/EU CLP Category 1 (JRC, 2004).	<p>A draft TG on IRE has never been submitted to OECD.</p> <p>The ESAC, 2007 released a statement "on the conclusions of the ICCVAM retrospective study on organotypic <i>in vitro</i> assays and recommended for the IRE test method, "that further work should be performed before a statement on [its] validity can be made" (ECVAM, 2007). However, ESAC emphasized that although not yet validated, positive outcomes from this test could be used as the basis for classifying and labelling chemicals as inducing serious eye damage (UN GHS/EU CLP Category 1) as</p>

Method	Test system	Description	Applicability	Regulatory status
				previously stated by European authorities (EC, 2004).
<i>Organotypic (chorio-allantoic membrane) methods</i>				
Hen's Egg Test - Chorioallantoic Membrane (HET-CAM)	CAM of a hen's egg	Permits the identification of irritant reactions which look similar to those occurring in the eye in the standard Draize rabbit eye test. In the HET-CAM assay, which is carried out with hen's eggs at the ninth day of embryonic development when the nervous system has not yet developed, reactions to test chemicals specifically haemorrhage, lysis, coagulation and to some extent hyperaemia of the chorio-allantoic membrane (CAM) are monitored after placing the test sample directly onto the CAM. A visual evaluation of the above mentioned parameters over a 5-minute observation period takes place.	Suitable for identifying at least UN GHS/EU CLP Category 1 (JRC, 2004). May also be suitable for identifying UN GHS/EU CLP No Category but more data are needed, especially for UN GHS/EU CLP Category 2 chemicals, to confirm this usefulness.	HET-CAM underwent formal validation by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for the UN GHS classification system, but was not considered useful at that time to be used for regulatory purposes for the evaluation of the eye irritation potential of chemicals (ICCVAM, 2006; ICCVAM, 2010). Its usefulness for identifying UN GHS/EU CLP No Category was not recommended by ICCVAM due to the lack of sufficient data for Category 2 chemicals (ICCVAM, 2010)

Method	Test system	Description	Applicability	Regulatory status
Reconstructed Human Tissue Models				
EpiOcular™ Eye Irritation Test (EIT)	EpiOcular™ Reconstructed human Tissue model	This tissue construct is a non-keratinized multilayered and stratified (but not cornified) epithelium prepared from normal human-derived epidermal keratinocytes. It is intended to model properties of the corneal epithelium. The reconstructed tissue is prepared in inserts with a porous membrane (MTI-003) through which the nutrients pass to the cells. A cell suspension is seeded into the MTI-003 membrane in specialized medium. After a period of initial cell proliferation, the medium is removed from the top of the tissue so that the epithelial surface is in direct contact with the air. This allows the test chemical to be directly applied to the epithelial surface, mimicking exposure of the corneal epithelium <i>in vivo/in situ</i> . This topical exposure is essential for modelling the same kind of progressive injury expected <i>in vivo</i> , allowing moreover the application of both solid and liquid chemicals. In the EpiOcular™ EIT, liquids and solids are applied using different exposure and post-exposure incubation times.		ESAC peer review will initiate Q4 2013
SkinEthic™ Human Corneal Epithelium (HCE)	SkinEthic™ HCE Reconstructed human Tissue model	The tissue construct consists of immortalized human corneal epithelial cells seeded on a polycarbonate membrane at the air-liquid interface. The tissue construct obtained is a multilayered epithelium resembling the <i>in vivo</i> corneal epithelium. As <i>in vivo</i> , columnar basal cells are present, including Wing cells. The model is characterized by the presence of ultra-structural features such as intermediate filaments, mature hemi-desmosomes and desmosomes that characterize the relevant epithelium <i>in situ</i> . Specific		Undergoing further optimisation

Method	Test system	Description	Applicability	Regulatory status
		cytokeratins 64kD (K.3) have also been described (Nguyen D.H. <i>et al.</i> , 2003).		
Cell-based assays				
Cytosensor Microphysiometer (CM)	Mouse L929 fibroblasts	Cytotoxicity and cell-function based <i>in vitro</i> assay that is performed on a sub-confluent monolayer of adherent mouse L929 fibroblasts cultured in a sensor chamber using a pH-meter to detect changes in acidity. Mechanistically, the CM test method is intended to model the cytotoxic action of an irritant chemical on the cell membranes of the cornea and conjunctival epithelium where the test chemical would reside in an <i>in vivo</i> exposure. The CM estimates the metabolic rate of a population of cells maintained in low volume flow-through chambers by measuring the rate of excretion of acid by-products and the resulting decrease in pH of the surrounding medium. The rate of change in pH per unit time becomes the metabolic rate of the population of cells. If a test chemical causes cytotoxicity to this population of cells it is assumed that the metabolic rate will fall. The CM may also address cell metabolism and recovery.	Suitable for identifying UN GHS/EU CLP Category 1 water-miscible chemicals (ECVAM, 2009) Suitable for identifying UN GHS/EU CLP No Category water-miscible surfactants and surfactant-containing mixtures (ECVAM, 2009).	Draft OECD TG (proposed 12/2012)
Fluorescein Leakage (FL)	Madin-Darby Canine Kidney (MDCK) cells	Toxic effects after a short exposure time of the test substance are measured by an increase in permeability of sodium fluorescein dye through the epithelial monolayer of Madin-Darby Canine Kidney (MDCK) cells cultured on permeable inserts. The amount of fluorescein leakage that occurs is proportional to the chemical-induced damage to the tight junctions, desmosomal junctions and cell membranes, and can be used to estimate the ocular toxicity potential of a test substance	Suitable for identifying UN GHS/EU CLP Category 1 water-miscible chemicals (ECVAM, 2009)	OECD TG 460 (adopted 10/2012)

Method	Test system	Description	Applicability	Regulatory status
Short Time Exposure (STE)	Statens Seruminstitut Rabbit Cornea (SIRC) cells	The STE test method is a cytotoxicity-based <i>in vitro</i> assay that is performed on a confluent monolayer of Statens Seruminstitut Rabbit Cornea (SIRC) cells, cultured on a 96-well polycarbonate microplate. The STE test method attempts to approximate the actual exposure in the human eye and evaluate cytotoxicity as an endpoint that measures the extent of damage to the SIRC cells following a five-minute exposure to a test chemical. In the STE test, cytotoxicity is quantitatively measured as the relative viability of SIRC cells. Cell viability is measured using an MTT assay (4). Decreased cell viability is predictive of adverse effects leading to corneal damage.	Suitable for identifying UN GHS/EU CLP Category 1 chemicals as well as UN GHS/EU CLP No Category surfactant solids and chemicals with a vapour pressure < 6 kPa (not highly volatile)	Draft OECD TG (proposed 11/2013)

Chapter 4. Table of References with Notes

OECD TG and draft updates relevant for Eye Irritation	
<p>OECD (2012). Test No. 405: Acute Eye Irritation/Corrosion, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070646-en</p>	<p>This Test Guideline (adopted in 1981 and updated in 1987, 2002, and 2012) includes the recommendation that prior to undertaking the described <i>in vivo</i> test for acute eye irritation/corrosion, a weight-of-the-evidence analysis be performed on the existing relevant data. Where insufficient data are available, it is recommended that they be developed through application of sequential testing. The testing strategy includes the performance of validated and accepted <i>in vitro</i> tests and is provided as a Supplement to the Guideline. Testing in animals should only be conducted if determined to be necessary after consideration of available alternative methods, and use of those determined to be appropriate. At the time of drafting of this updated TG 405, there are instances where using this Test Guideline is still necessary or required by some regulatory authorities.</p> <p>The latest update (2012) mainly focused on the use of analgesics and anesthetics without impacting the basic concept and structure of the Test Guideline. The refinements described in this proposal will substantially reduce or avoid animal pain and distress in most testing situations where <i>in vivo</i> ocular safety testing is still necessary.</p>
<p>OECD (2013). Test No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076303-en</p>	<p>The Bovine Corneal Opacity and Permeability test method (BCOP) is an <i>in vitro</i> test method that can be used to identify chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) as well as chemicals not requiring classification for serious eye damage/eye irritation. The BCOP uses isolated corneas from the eyes of cattle slaughtered for commercial purposes, thus avoiding the use of laboratory animals. Each treatment group (test substance, negative/positive controls) consists of a minimum of three eyes where the cornea has been excised and mounted to a holder. Depending on the physical nature and chemical characteristics of the test substance, different methods can be used for its application since the critical factor is ensuring that the test substance adequately covers the epithelial surface.</p> <p>While it is not considered valid as a complete replacement for the <i>in vivo</i> rabbit eye test, the BCOP is recommended for use as part of a tiered-testing strategy for regulatory classification and labelling within a specific applicability domain. Test</p>

	<p>chemicals (including substances and mixtures) can be classified/identified as UN GHS/EU CLP Category 1 or No Category without further testing in rabbits.</p> <p>An update of the TG 437 took place in 2013. The main differences between the original 2009 version and the 2013 updated version concern, but are not limited to: the use of the BCOP test method to identify chemicals not requiring classification according to UN GHS (paragraphs 2 and 7); clarifications on the applicability of the BCOP test method to the testing of alcohols, ketones and solids (paragraphs 6 and 7) and of substances and mixtures (paragraph 8); clarifications on how surfactant substances and surfactant-containing mixtures should be tested (paragraph 28); updates and clarifications regarding the positive controls (paragraphs 39 and 40); an update of the BCOP test method decision criteria (paragraph 47); an update of the study acceptance criteria (paragraph 48); an update to the test report elements (paragraph 49); an update of Annex 1 on definitions; the addition of Annex 2 for the predictive capacity of the BCOP test method under various classification systems; an update of Annex 3 on the list of proficiency chemicals; and an update of Annex 4 on the BCOP corneal holder (paragraph 1) and on the opacitometer (paragraphs 2 and 3).</p>
<p>OECD (2013). Test No. 438: Isolated Chicken Eye Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076310-en</p>	<p>The Isolated Chicken Eye (ICE) test method is an <i>in vitro</i> test method that can be used to identify chemicals inducing serious eye damage (UN GHS/EU CLP Category 1) as well as chemicals not requiring classification for serious eye damage/eye irritation. The ICE method uses eyes collected from chickens obtained from slaughterhouses where they are killed for human consumption, thus eliminating the need for laboratory animals. The eye is enucleated and mounted in an eye holder with the cornea positioned horizontally. The test chemical and negative/positive controls are applied to the cornea. Toxic effects to the cornea are measured by a qualitative assessment of opacity, a qualitative assessment of damage to epithelium based on fluorescein retention, a quantitative measurement of increased thickness (swelling), and a qualitative evaluation of macroscopic morphological damage to the surface. The endpoints are evaluated separately to generate an ICE class for each endpoint, which are then combined to generate an Irritancy Classification for each test substance.</p> <p>An update of the TG 438 took place in 2013. The main differences between the original 2009 version and the 2013 updated version concern, but are not limited</p>

	to: the use of the ICE test method to identify chemicals not requiring classification according to UN GHS Classification System, an update to the test report elements, an update of Annex 1 on definitions, and an update to Annex 2 on the proficiency chemicals.
OECD (2012). Test No. 460: Fluorescein Leakage Test Method for Identifying Ocular Corrosives and Severe Irritants, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264185401-en	<p>This Test Guideline describes an <i>in vitro</i> assay that may be used for identifying water soluble ocular corrosives and severe irritants as defined by the UN Globally Harmonized System of Classification and Labelling, Category 1. The assay is performed in a well where a confluent monolayer of Madin-Darby Canine Kidney (MDCK) is used as a separation between two chambers. It uses a fluorescein dye as marker. The test substance has the potential to impair the junctions of the MDCK cells and thus to increase the monolayer's permeability. Consequently the fluorescein passes through the monolayer and the fluorescein leakage (FL) increases. The FL is calculated as a percentage of leakage relative to both a blank control and a maximum leakage control. The concentration of test substance that causes 20% FL (FL₂₀, in mg/mL) is calculated and used in the prediction model for identification of ocular corrosive and severe irritants. The cut-off value of FL₂₀ to identify water soluble chemicals as ocular corrosives/severe irritants is ≤100mg/mL. The FL test method should be part of testing strategies in which strategic combinations of several alternative test methods may be able to replace the <i>in vivo</i> eye test. The Top-Down approach (Scott <i>et al.</i> 2010) is designed to be used when, based on existing information, a chemical is expected to have high irritancy potential.</p> <p>The FL test method can identify substances within a limited applicability domain as ocular corrosives/severe irritants (UN GHS Category 1; EU CLP Category 1; U.S. EPA Category I) without any further testing. The same is assumed for mixtures although mixtures were not used in the validation. Therefore, the FL test method may be used to determine the eye irritancy/corrosivity of chemical following the sequential testing strategy of TG 405.</p>
<i>In vitro</i> test methods	
Pfannenbecker U, Bessou-Touya S, Faller C, et al. (2013): Cosmetics Europe multi-laboratory pre-validation of the EpiOcular™ reconstituted human tissue test method for the prediction of eye irritation. <i>Toxicol In Vitro</i> 27: 619-626.	Cosmetics Europe, The Personal Care Association (known as Colipa before 2012), conducted a program of technology transfer and within/between laboratory reproducibility of MatTek Corporation's EpiOcular™ Eye Irritation Test (EIT) as one of the two human reconstructed tissue test methods. This this test method uses a single exposure period for each chemical and a prediction model based on a cut-off in relative survival [≤60% = irritant (I) (UN GHS/EU CLP Categories 2

	<p>and 1); >60% = no classification (UN GHS/EU CLP No Category)]. Two different protocols with different exposure times and post-treatment incubations are used for liquid and solid chemicals. Tissue viability is determined by tetrazolium dye (MTT) reduction. In the Cosmetics Europe study, combinations of 20 coded chemicals were tested in 7 laboratories. Standardized laboratory documentation was used by all laboratories. Twenty liquids (11 NC/9 I) plus 5 solids (3 NC/2 I) were selected so that both exposure regimens could be assessed. Concurrent positive (methyl acetate) and negative (water) controls were tested in each trial. In all, 298 independent trials were performed and demonstrated 99.7% agreement in prediction (NC/I) across the laboratories. Coefficients of variation for the % survival for tissues from each treatment group across laboratories were generally low. This test method entered a formal ECVAM validation study in December 2008.</p>
<p>Alépée N, Bessou-Touya S, Cotovio J, De Smedt A, De Wever B, Faller C, Jones P, Le Varlet B, Marrec-Fairley M, Pfannenbecker U, Tailhardat M, Van Goethem F & McNamee P (2013). Cosmetics Europe Multi-Laboratory Pre-Validation of the SkinEthic™ Reconstituted Human Corneal Epithelium Test Method for the Prediction of Eye Irritation. <i>Toxicology in Vitro</i> 27: 1476-1488.</p>	<p>Cosmetics Europe, The Personal Care Association, known as Colipa before 2012, conducted a program of technology transfer and assessment of Within/Between Laboratory (WLV/BLV) reproducibility of the SkinEthic™ Reconstituted Human Corneal Epithelium (HCE) as one of two human reconstructed tissue eye irritation test methods. The SkinEthic™ HCE test method involves two exposure time treatment procedures – one for short time exposure (10 min – SE) and the other for long time exposure (60 min – LE) of tissues to test substance. This paper describes pre-validation studies of the SkinEthic™ HCE test method (SE and LE protocols) as well as the Eye Peptide Reactivity Assay (EPRA). In the SE WLV study, 30 substances were evaluated. A consistent outcome with respect to viability measurement across all runs was observed with all substances showing an SD of less than 18%. In the LE WLV study, 44 out of 45 substances were consistently classified. These data demonstrated a high level of reproducibility within laboratory for both the SE and LE treatment procedures. For the LE BLV, 19 out of 20 substances were consistently classified between the three laboratories, again demonstrating a high level of reproducibility between laboratories. The results for EPRA WLV and BLV studies demonstrated that all substances analysed were categorised similarly and that the method is reproducible. The SkinEthic™ HCE test method entered a formal ECVAM validation study in 2010.</p>
<p>Van Goethem F, Adriaens E, Alépée N, Straube F, De Wever B, Cappadoro M, Catoire S, Hansen E, Wolf A & Vanparys P (2006).</p>	<p>This multicentre study aimed at evaluating the reliability (reproducibility) and relevance (predictivity) of a new commercially available human corneal epithelial</p>

<p>Prevalidation of a new <i>in vitro</i> reconstituted human cornea model to assess the eye irritating potential of chemicals. <i>Toxicology in Vitro</i> 20: 1-17.</p>	<p>(HCE) model (SkinEthic Laboratories, Nice, France) to assess acute ocular irritation. A prevalidation approach (protocol optimization, transfer and performance) was followed and at each of the four participating laboratories, 20 coded reference chemicals, covering the whole range of irritancy, were tested. The compounds were applied topically to the HCE cultures and the level of cytotoxicity (tissue viability and histological analysis) was determined. Once a standardised protocol was established, a high level of reproducibility between the laboratories was observed. In order to assess the capability of the HCE model to discriminate between irritants (I) and non-irritants (NI), a classification prediction model (PM) was defined based on a viability cut-off value of 60%. The obtained <i>in vitro</i> classifications were compared with different <i>in vivo</i> classifications (e.g. Globally Harmonised System) which were calculated from individual rabbit data described in the ECETOC data bank. Although an overall concordance of 80% was obtained (sensitivity = 100% and specificity = 56%), the predictivity of the HCE model substantially increased when other sources of <i>in vivo</i> and <i>in vitro</i> data were taken into account.</p> <p>NB This reference is now dated, since formal validations studies have taken place since then (see main text).</p>
<p>QSAR</p>	
<p>Solimeo R, Zhang J, Kim M, Sedykh A & Zhu H (2012). Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach. <i>Chemical Research in Toxicology</i> 25 (12): 2763-2769</p>	<p>These authors have developed quantitative structure-activity relationship (QSAR) models for a set of small molecules with animal ocular toxicity data compiled by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods. The data set was initially curated by removing duplicates, mixtures, and inorganics. The remaining 75 compounds were used to develop QSAR models. Both k nearest neighbor and random forest statistical approaches in combination with Dragon and Molecular Operating Environment descriptors were applied. Developed models were validated on an external set of 34 compounds collected from additional sources. The external correct classification rates (CCR) of all individual models were between 72 and 87%. Furthermore, the consensus model, based on the prediction average of individual models, showed additional improvement (CCR = 0.93). The validated models could be used to screen external chemical libraries and prioritize chemicals for <i>in vivo</i> screening as potential ocular toxicants.</p>

<p>Gallegos Saliner A, Patlewicz G & Worth AP (2006). Review of Literature-Based Models for Skin and Eye Irritation and Corrosion. JRC Report EUR 22320 EN. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/QSAR_Review_Irritation.pdf</p>	<p>This report reviews the state-of-the-art of <i>in silico</i> and <i>in vitro</i> methods for assessing dermal and ocular irritation and corrosion. Following a general introduction, the current EU legislation for the classification and labelling of chemicals causing irritation and corrosivity is summarized. Then currently available non-animal approaches are reviewed. The main alternative approaches to assess acute local toxic effects are: a) <i>in silico</i> approaches, including SARs, QSARs and expert systems integrating multiple approaches; and b) <i>in vitro</i> test methods. In this review, emphasis is placed on literature-based (Q)SAR models for skin and eye irritation and corrosion as well as computer-based expert systems.</p>
<p>Gallegos Saliner A, Patlewicz G & Worth AP (2008). A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. QSAR & Combinatorial Science 27: 49-59.</p>	<p>This paper reviews the state-of-the-art of <i>in silico</i> methods for assessing dermal and ocular irritation and corrosion. It is based on an in-depth review performed by the European Chemicals Bureau of the European Commission: Joint Research Centre in support of the development of technical guidance for the implementation of the REACH legislation, and is one of a series of minireviews in this journal. The most widely used <i>in silico</i> approaches are classified into methods to assess (1) skin irritation, (2) skin corrosion and (3) eye irritation. In this review, emphasis is placed on literature-based (Q)SAR models.</p>
<p>Tsakovska I, Netzeva T & Worth AP (2005). Evaluation of (Q)SARs for the prediction of Eye Irritation/Corrosion Potential - physicochemical exclusion rules. European Commission Report EUR 21897 EN. European Commission-Joint Research Centre, Ispra, Italy. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology</p>	<p>A set of rules (rulebase) has been developed by the German Bundesinstitut für Risikobewertung (BfR) and incorporated into a Decision Support System (DSS), which is used within the BfR to support regulatory decisions. The rulebase is based on the combined use of two predictive approaches: a) physicochemical exclusion rules to identify chemicals with no skin irritation/corrosion or eye irritation/corrosion potential; and b) structural inclusion rules to identify chemicals with skin irritation/corrosion or eye irritation/corrosion potential.</p> <p>In this study, an evaluation was performed of the physicochemical rulebase (comprising 31 physicochemical exclusion rules) for predicting the absence of eye irritation/corrosion. In particular, the following aspects were addressed: a) an assessment of the derivation of the rules by using the rulebase training set of 1358 substances (343 irritants/corrosives and 1015 non-irritants and non-corrosives); b) an external validation by using a test set of 199 chemicals (45 irritants/corrosives and 154 non-irritants and non-corrosives); c) an assessment of the suitability of the test set</p>

	<p>used for validation; d) an assessment of the compliance of the rule-base with the OECD principles for the validation of (Q)SARs. According to the results of this study: a) the physicochemical exclusion rules for eye irritation/corrosion comply well with the OECD validation principles; b) predictions of no adverse effect (NOT R34/R35/R36/R41) can be made for 20 out of the 199 chemicals in the test set; c) 3 of the 45 irritants/corrosives are falsely predicted as non-irritant or non corrosive; d) the probability of a negative prediction being correct (Negative Predictive Value) is 0.87; and e) approximately 10% of Draize rabbit eye tests could be avoided by relying on the predictions of no adverse effect. On the basis of the results, some recommendations were made regarding the improvement of the physicochemical exclusion rules for eye irritation/corrosion.</p>
<p>Tsakovska I, Gallegos A, Netzeva T, Pavan M & Worth AP (2007). Evaluation of SARs for the prediction of eye irritation/corrosion potential – structural inclusion rules in the BfR decision support system. SAR and QSAR in Environmental Research 18, 221-235.</p>	<p>The proposed REACH regulation within the European Union (EU) aims to minimise the number of laboratory animals used for human hazard and risk assessment while ensuring adequate protection of human health and the environment. One way to achieve this goal is to develop non-testing methods, such as (quantitative) structure–activity relationships ([Q]SARs), suitable for identifying toxicological hazard from chemical structure and physicochemical properties alone. A database containing data submitted within the EU New Chemicals Notification procedure was compiled by the German Bundesinstitut für Risikobewertung (BfR). On the basis of these data, the BfR built a decision support system (DSS) for the prediction of several toxicological endpoints. For the prediction of eye irritation and corrosion potential, the DSS contains 31 physicochemical exclusion rules evaluated previously by the European Chemicals Bureau (ECB), and 27 inclusion rules that define structural alerts potentially responsible for eye irritation and/or corrosion. This work summarises the results of a study carried out by the ECB to assess the performance of the BfR structural rulebase. The assessment included: (a) evaluation of the structural alerts by using the training set of 1341 substances with experimental data for eye irritation and corrosion; and (b) external validation by using an independent test set of 199 chemicals. The test set of 199 substances contained 154 (77%) non-labeled substances and 45 (23%) labeled as eye irritants/corrosives, subdivided as follows: (i) 10 R36 substances (5%); (ii) 28 R41 substances (14%); and (iii) 7 substances (4%) labeled R34 or R35. Recommendations are made for the further</p>

	development of the structural rules in order to increase the overall predictivity of the DSS.
Abraham MH, Hassanisadi M, Jalali-Heravi M, Ghafourian T, Cain WS & Cometto-Muniz JE (2003). Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis. <i>Toxicological Sciences</i> 76(2):384-391.	Draize rabbit eye test scores, as modified maximum average score (MMAS), for 68 pure bulk liquids were adjusted by the liquid-saturated vapor pressure P. These 68 adjusted scores, as $\log(\text{MMAS}/P)$, were shown to be completely equivalent to eye irritation thresholds (EIT), expressed as $\log(1/\text{EIT})$, for 23 compounds in humans. Thus, for the first time the Draize eye test in rabbits for pure bulk liquids is shown to be perfectly compatible with eye irritation thresholds in humans. The total data set for 91 compounds was analyzed by the general solvation equation of Abraham. Values of $\log(\text{MMAS}/P)$ or $\log(1/\text{EIT})$ could be fitted to a five-parameter equation with $R^2 = 0.936$, $SD = 0.433$, $AD = 0.000$, and $AAD = 0.340$ over a range of 9.6 log units. When divided into a training set of 45 compounds, the corresponding equation could be used to predict the remaining 46 compounds in a test set with $AD = -0.037$ and $AAD = 0.345$ log units. Thus, the 91-compound equation can now be used to predict further EIT values to around 0.4 log units. It is suggested that the mechanism of action in the Draize test and in the human EIT involves passive transfer of the compound to a biophase that is quite polar, is a strong hydrogen bond base, a moderate hydrogen bond acid, and quite hydrophobic. The biophase does not resemble water or plasma, but resembles an organic solvent such as N-methylformamide.
Testing Strategies	
Gallegos Saliner A & Worth AP (2007). Testing Strategies for the Prediction of Skin and Eye Irritation and Corrosion for Regulatory Purposes. JRC Report EUR 22881 EN. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/	This report reviews the use of stepwise testing approaches for the prediction of skin and eye irritation and corrosion in a regulatory context. It is published as a companion report to the Review of Literature-Based Models for Skin and Eye Irritation and Corrosion, an ECB report which reviewed the state-of-the-art of <i>in silico</i> and <i>in vitro</i> dermal and ocular irritation and corrosion human health hazard endpoints. In the former review, the focus was placed on reviewing alternative <i>in silico</i> approaches to assess acute local toxic effects, such as QSARs, SARs, chemical categories, and read-across and analogue approaches. Special emphasis was placed on literature-based (Q)SAR models for skin and eye irritation and corrosion and expert systems. In the present review, the emphasis is on different schemes (testing strategies) that have been conceived for the integrated use of different approaches, including <i>in silico</i> , <i>in vitro</i> and <i>in vivo</i> methods.
Grindon C, Combes R, Cronin MT, Roberts DW, Garrod JF (2008). An integrated decision-tree testing strategy for eye irritation with	This paper presents some results of a joint research project, sponsored by Defra and conducted by FRAME and Liverpool John Moores University, on the status

<p>respect to the requirements of the EU REACH legislation. ATLA 6(1):81-92.</p>	<p>of alternatives to animal testing with regard to the European Union REACH (Registration, Evaluation and Authorisation of Chemicals) system for the safety testing and risk assessment of chemicals. The project covered all the main toxicity endpoints associated with REACH. This paper focuses on the use of alternative (non-animal) methods (both <i>in vitro</i> and <i>in silico</i>) for eye irritation testing. The manuscript reviews numerous <i>in vitro</i> tests and their possible collation into test batteries, <i>in silico</i> models and a refined <i>in vivo</i> method (the low volume eye test), before combining the use of all these methods into an integrated testing strategy. The aim of this strategy is a reduction in the number animal tests which would need to be performed in the process of fulfilling the REACH system criteria; this would also lead to a lowering of the number of animals required in compliance with the REACH system requirements.</p>
<p>Scott L, Eskes C, Hoffmann S, et al. (2010) A proposed eye irritation testing strategy to reduce and replace <i>in vivo</i> studies using Bottom-Up and Top-Down approaches. <i>Toxicology in Vitro</i> 24: 1-9.</p>	<p>This report shows the outcome of an Expert Meeting convened by ECVAM in February 2005 to identify test strategies for eye irritation. In this workshop test developers/users were requested to nominate methods to be considered as “building blocks” to populate such testing strategies. Assays were evaluated and categorized based on their proposed applicability domains (e.g., categories of irritation severity, modes of action, chemical class, and physicochemical compatibility). The difficulty in predicting the middle category of irritancy (e.g. R36, GHS Categories 2A and 2B) was recognized. The testing scheme proposes using a Bottom-Up (begin with using test methods that can accurately identify non-irritants) or Top-Down (begin with using test methods that can accurately identify severe irritants) progression of <i>in vitro</i> tests (based on expected irritancy). Irrespective of the starting point, the approach would identify non-irritants and severe irritants, leaving all others to the (mild/moderate) irritant GHS 2/R36 categories. Since then, all <i>in vitro</i> test methods have been validated and were regulatory adopted within the top-down/bottom-up framework.</p>

5. Skin sensitisation

Silva Casati, Coralie Dumont & Andrew Worth

5.1 *The traditional animal test and its regulatory use*

Allergic contact dermatitis (ACD), the clinical manifestation of skin sensitisation, is the result of occupational, environmental or domestic exposure to sensitising chemicals and represents the most frequent manifestation of immunotoxicity in humans.

Skin sensitisation is an immunological process which develops in two distinct phases. The first phase, *induction*, is initiated when a susceptible individual is exposed to a contact allergen in a sufficient amount to trigger the proliferation of memory T lymphocytes. During the second phase, *elicitation*, which occurs following a subsequent contact with the allergen, the expanded population of memory T lymphocytes reacts to the sensitising agent, leading to the allergic response with the clinical signs of ACD.

The regulatory assessment of skin sensitisation has traditionally involved the use of animals. These include the classical methods based on guinea-pigs: the Magnusson Kligman Guinea Pig Maximisation Test (GPMT; OECD, 1992), an adjuvant-type test in which the allergic response is potentiated by intradermal injection of Freund's Complete Adjuvant (FCA), and the Buehler test (OECD, 1992), a non-adjuvant technique that involves topical induction treatment only. More recently, a murine test, the Local Lymph Node Assay (LLNA; OECD, 2010a), has gained acceptance as an alternative to the guinea-pig tests, with the advantage of providing both refinement and reduction of animal usage and a more objective and quantitative endpoint.

The mouse and guinea-pig methods differ with respect to the endpoints evaluated. The LLNA measures the responses provoked during the induction of sensitisation, and specifically the extent of proliferation of lymphocytes in regional lymph nodes draining the site of application of the test substance measured through quantification of radioactive label incorporation (Kimber & Basketter, 1992). Non-radioactive variations of the LLNA have also been developed (OECD, 2010b,c). The two guinea-pig tests measure elicitation reactions in the skin in previously sensitised animals. The GPMT is considered to be more sensitive than the Buehler assay. The LLNA generates dose-response information from which an EC₃ value (concentration of substance needed to elicit a positive response at the threshold level) can be derived. The EC₃ value can be used as an indicator of relative potency in risk assessment, or to classify chemicals on the basis of potency bands.

Information on skin sensitisation (potential or potency) is required to support the human safety assessment of substances under several pieces of EU legislation. Information on sensitisation potential is needed for the hazard assessment of industrial chemicals (European Commission, 2006), biocides (European Union, 2012) plant protection products (European Commission, 2009a) and for classification and labelling (European Commission, 2008). For cosmetics, information on sensitisation potency is additionally required, and this must be obtained without recourse to animal testing, since testing and marketing bans are in force (European Commission, 2009b).

5.2 Mechanistic understanding of the endpoint

The biological mechanisms underlying the induction and elicitation phases of skin sensitisation (Figures 5.1 and 5.2) are relatively well understood and extensively documented in the scientific literature (Vocanson et al., 2009; Basketter & Kimber, 2010).

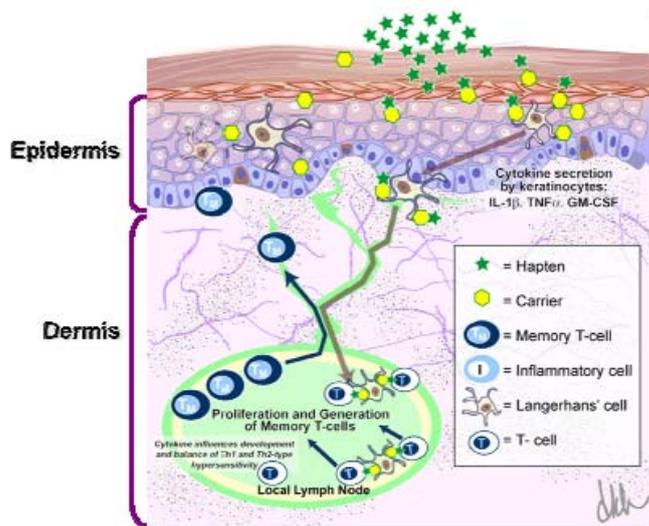


Figure 5.1. The induction phase of skin sensitisation (from OECD 2012a)

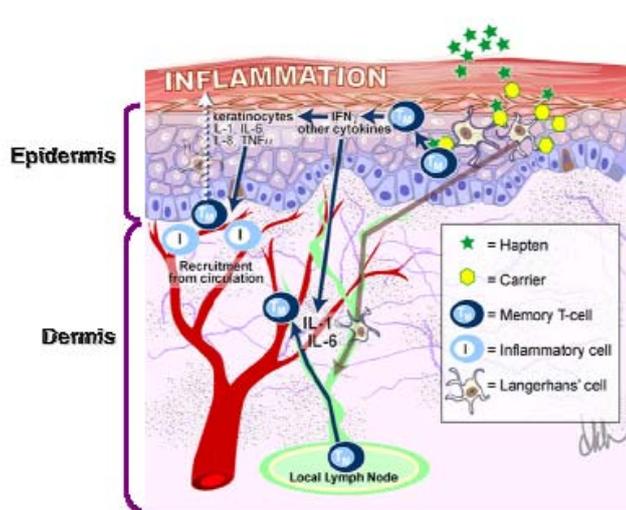


Figure 5.2. The elicitation phase of skin sensitisation (from OECD 2012a)

The key events in the skin sensitisation pathway include: 1) the ability of the substance to cross the stratum corneum of the skin and reach the viable epidermis (epidermal disposition or bioavailability); 2) the covalent binding of the substance (hapten) to skin proteins (haptentation); 3) the release of pro-inflammatory cytokines and the induction of cytoprotective cellular pathways in keratinocytes (KC); 4) the activation and maturation of dendritic cells (DC), the skin's immunocompetent cells, and their migration from skin to the regional lymph nodes; 5) the presentation by DCs of the haptened protein to T cells (T-cell priming) and the clonal expansion of memory T cells (Adler et al., 2011). While there is general agreement on these key events, a detailed understanding of the underlying chemical and biological interactions remains incomplete. More recently, the existing knowledge has

been recast in form of an Adverse Outcome Pathway (AOP) by the OECD (OECD 2012a,b; Figure 5.3).

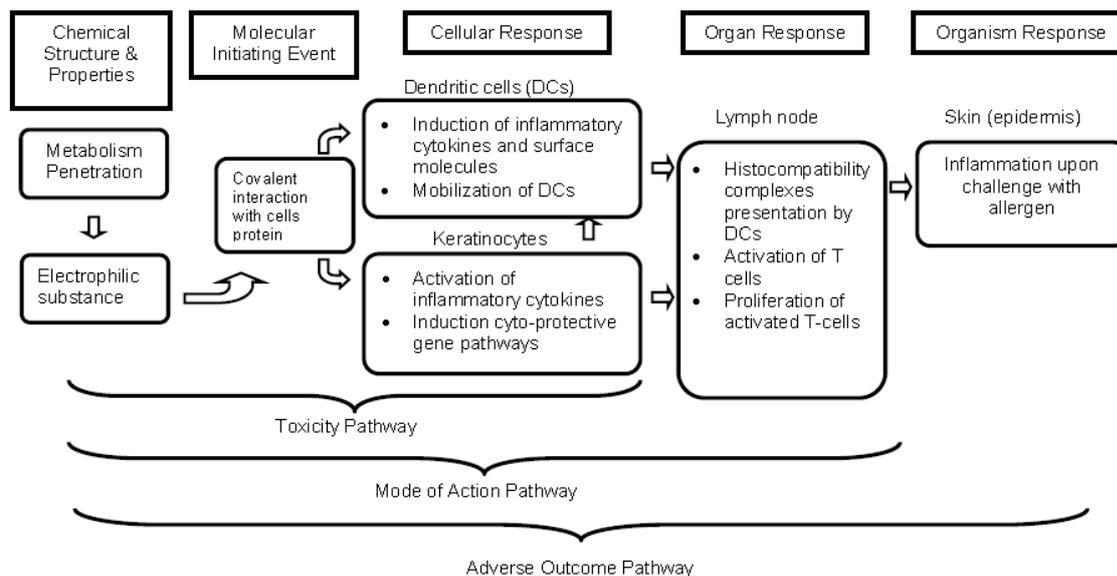


Figure 5.3. Adverse Outcome Pathway for Skin Sensitisation (OECD, 2012a)

The majority of skin sensitising substances share the common feature of being low molecular weight reactive chemicals that are not able to elicit an immune response on their own but need to bind to proteins to form a hapten-protein conjugate, which is recognised by the immune system. Reactivity towards skin proteins is considered to be the main Molecular Initiating Event (MIE) needed to trigger the cascade of subsequent events. While organic chemicals bind proteins through stable covalent bonds, metals can form coordination complexes.

A proportion of the organic skin sensitisers are not directly reactive but need to be activated either by cutaneous enzymes (pro-haptens) or by abiotic transformation (pre-haptens) (e.g. by air oxidation) (Lepoittevin, 2006). There is little information on the role of human skin metabolism in the sensitisation pathway.

5.3 Status of non-standard methods and Integrated Testing Strategies

Our understanding of the chemistry and biology of skin sensitisation has favoured the development in recent years of alternative test methods designed to detect and measure the key events in the skin sensitisation pathway (Adler et al., 2011).

Penetration through the stratum corneum to the viable epidermis is essential for a chemical to exert sensitisation activity. While there are no officially accepted models for epidermal disposition (Basketter et al., 2007), some physicochemical properties of the compound can be used to provide information about its potential to penetrate the skin. The molecular weight (MW), the octanol-water partition coefficient ($\log K_{OW}$ or $\log P$) and the ionisation state are considered to be useful physicochemical predictors (World Health Organization (WHO) 2006; Roberts and Williams, 1982). $\log K_{OW}$ is commonly used in Quantitative Structure Activity Relationship (QSAR) models. A common shortcoming of available QSARs is that they do not account for solvent and formulation effects, and are typically based on the assumption of passive diffusion through a homogeneous compartment. In addition,

mathematically-based and more biologically faithful toxicokinetic models have been developed to simulate the skin penetration process (e.g. Dancik et al., 2013).

The ability to detect pre-haptens and pro-haptens remains a challenge, since these sensitisers are not systematically detected by test methods that lack or have limited metabolic and oxidative capacity. However, expert systems such as TIMES-SS (Patlewicz et al., 2007; Roberts et al., 2007) and the OECD QSAR toolbox (<http://www.qsartoolbox.org>) provide a means of simulating abiotic and enzymatic reactions and incorporating this information into their predictions.

The recognition that reactivity towards skin proteins is the main MIE in the skin sensitisation pathway has led to the development of *in chemico* methods (based on organic chemistry) to detect and quantify chemical reactivity. The majority of these methods monitor either the depletion of a nucleophile and/or the formation of an adduct between an electrophilic chemical and the nucleophile. The experimental nucleophile is an organic compound (such as nitrobenzenethiol) or a peptide like cysteine, lysine and glutathione (GSH) (Gerberick et al., 2008). Some of these assays have been adapted to provide information on reaction kinetics (Schultz et al., 2005; Böhme et al., 2009). These approaches are generally limited by the lack of metabolism and oxidation capacity, although a recent development of one of these assays incorporates an-enzyme mediated activation step for the identification of pro-haptens (Gerberick et al., 2009). The availability of *in chemico* data has enabled the development of QSARs for protein reactivity, as reviewed by Asturiol & Worth (2011).

KCs are the predominant cells in the skin and play a role in the initiation of the immune response by releasing a wide range of pro-inflammatory mediators and growth factors (Griffiths et al., 2005). Measurement of intracellular interleukin-18 (IL-18) production in a human keratinocyte cell line following exposure to sensitisers has been proposed as a biomarker for the detection of skin sensitisers and for discriminating between skin and respiratory sensitisers (Galbiati et al., 2011). Studies in animals have shown that the antioxidant/electrophile response element ARE/EpRE-dependent pathway, which plays a central role in cellular defence against oxidative and electrophilic stress, is relevant to the skin sensitisation process (Kim et al., 2008). Accordingly, endpoints associated with the activation of this pathway are used in several *in vitro* methods based on keratinocyte reporter cell lines (Emter et al., 2010; Bauch et al., 2012) or the 3-dimensional reconstituted human epidermis (RhE) (McKim et al., 2010).

The next key event in the skin sensitisation pathway is the recognition and internalisation of the hapten-protein conjugate by immature DCs. During this process, DCs undergo a process of maturation in which they lose their antigen processing capacity and acquire the capability to present the allergen to naïve T lymphocytes. DC maturation is typically measured experimentally by assessing the expression of co-stimulatory and intercellular adhesion molecules (e.g. CD40, CD54, CD83, CD86) on the surface of DCs that are needed for the interaction with, and activation of T lymphocytes, or by quantifying cytokine secretion (e.g. IL-1 β , IL-6, IL-8) (Casati et al., 2005). Several test methods based on the measurement of biomarkers for DC activation are available (Adler et al., 2011), including DCs derived from peripheral blood, cord blood or bone-marrow samples, as well as DC-like cell lines (e.g. THP-1, U-937, MUTZ 3). Genomic and proteomic signatures in KC-based and DC-based assays have also been proposed as predictive markers for both skin sensitisation potential and potency assessment (Johansson et al., 2013; Lambrechts et al., 2010).

Following maturation, DCs migrate from the skin to regional lymph nodes, where they present the antigen to responsive T-lymphocytes. The activation and clonal expansion of naïve T-lymphocytes represent the last key event in the skin sensitisation pathway, and is the endpoint measured in the LLNA. The majority of T-cell based assays measure the *in vitro* proliferation of naïve T-cells (isolated from the peripheral blood mononuclear cells of healthy donors), following co-culture with sensitiser-treated DCs (Martin et al., 2010). Although these methods may represent the most specific tool for the identification of contact allergens, they involve rather complex DC-T cell co-culture procedures which have made it difficult to standardise the protocols. Therefore the availability of a robust test method based on *in vitro* T-cell proliferation will require further development.

Most QSARs and expert systems have been developed to be directly predictive of the sensitisation response in animals, as reviewed elsewhere (Patlewicz & Worth, 2008; Chaudhry et al, 2010; Adler et al., 2011). Some of these approaches (e.g. Toxtree and Derek) are based on mechanistic/expert knowledge, whereas others are based on statistical correlations (e.g. CAESAR, TOPKAT, MultCASE), or are best characterised as hybrid (based on both approaches, e.g. TIMES-SS). For the quantitative prediction of potency, mathematically-based (toxicokinetic and toxicodynamic) models (e.g. Maxwell & MacKay, 2008) will be needed. These models require a considerable amount of experimental data for calibration.

In order to replace animal testing in the assessment of skin sensitisation, a combination of different alternative methods addressing the key events in the skin sensitisation pathway will be needed. Test methods under formal validation by EURL ECVAM are being assessed for their reliability, with a view to their possible use within integrated assessment approaches. These methods include: 1) the Direct Peptide Reactivity Assay (DPRA) which addresses protein binding by monitoring the depletion of a nucleophile-containing synthetic peptide (Gerberick et al., 2007); 2) the KeratinoSens assay which measures the activity of the antioxidant/electrophile response element ARE/EpRE-dependent pathway in keratinocytes (Emter et al., 2010); and 3) the human Cell Line Activation Test (h-CLAT) which measures the induction of CD54 and CD86 protein markers on the surface of THP-1 cell lines (Sakaguchi et al., 2006; Ashikaga et al., 2010). These methods appear sufficiently reproducible to justify their inclusion in integrated approaches for hazard identification and classification. Their potential to contribute to the potency assessment of cosmetic ingredients is being evaluated by the Cosmetics Europe Skin Tolerance Task Force.

Following an initial proposal for integrating information from alternative approaches by Jowsey et al. (2006), a number of different approaches for data integration have emerged. These range from sequential testing strategies (STS), in which the decision making follows a stepwise approach involving interim decision steps (Nukada et al., 2013; van der Veen et al., 2014), weight-of-evidence based approaches (e.g. Natsch et al., 2009; Bauch et al., 2012) to mathematically based approaches such as Bayesian Networks, which encode probabilistic relationships among the input variables (Jaworska et al., 2011; 2013), and artificial neural networks (Hirota et al., 2013). These approaches are all mechanistically-based, in the sense that they use predictor variables associated with Key Events in the established AOP, and they are all data-driven, in the sense that statistical methods are used to optimise the models. An advantage of the weight of evidence approach is its simplicity, whereas an advantage of the probabilistic approach is that it provides information on the interdependences of all variables and is optimised to reduce the overall uncertainty of prediction.

In addition, the use of mathematical models to represent the mechanisms underlying skin sensitisation induction is being explored (Maxwell & MacKay, 2008; Maxwell et al., 2014).

EURL ECVAM (EURL ECVAM, 2013) has initiated the development of non-animal testing strategies for hazard identification and potency sub-categorisation according to the two subcategories 1A and 1B of the Globally Harmonised System of Classification and Labelling of Chemicals (GHS). This will contribute to a new OECD activity aimed at developing generic guidance on how to document and evaluate Integrated Approaches to Testing and Assessment (IATA) for skin sensitisation. EURL ECVAM is developing an IATA to meet the needs of REACH and CLP, with a view to reducing the need for animal testing in relation to the next registration deadline of 2018.

5.4 Conclusions

Information on skin sensitisation potential is required for hazard identification under REACH, the Biocides Directive, the Plant Protection Products Regulation, and can be used for classification and labelling under CLP. For cosmetics, additional information on sensitisation potency is required.

The mechanisms underlying skin sensitisation are sufficiently well understood to enable the development of non-standard methods addressing the key events of the skin sensitisation pathway. *In vitro* methods designed to cover haptentation, keratinocyte responses and DC activation are well developed, while further development efforts are needed to make robust T-cell proliferation assays available. Some of these methods are in an advanced status of formal validation by EURL ECVAM, and the development of corresponding OECD test guidelines has started.

QSAR models designed to predict key events in the sensitisation pathway should be useful sources of information in IATA, whereas QSARs that are directly predictive of the sensitisation response may be useful for increasing confidence in Weight of Evidence (WoE) arguments. QSARs are available for predicting skin penetration and protein binding, as well as the adverse outcome in animals, but models for the intermediate events are lacking. Mathematically based (systems biology) models provide the most promising means of simulating the dynamics of skin sensitisation and predicting potency, but these approaches need to be further developed and implemented in user-friendly tools.

Efforts are underway to develop integrated approaches for predicting skin sensitisation potential and potency based on the use of *in silico*, *in chemico*, and *in vitro* methods. AOP-based IATA, which are motivated by an understanding of the underlying mechanistic pathways, will be particularly useful for regulatory purposes. With a view to promoting the international acceptance of these integrated approaches, the OECD will develop guidance on how to document and evaluate IATA in view of their intended applications. The development of a single IATA, suitable for all chemicals and regulatory applications, is unrealistic, but a systematic approach to the description and evaluation of IATA will help the user to make an appropriate choice.

The extent and quality of the non-standard information needed in a regulatory assessment will depend on the regulatory purpose and the level of confidence required. In general, a greater burden of proof will be required to conclude on the absence of skin sensitisation potential than to conclude on its presence.

5.5 References

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485

Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N & Itagaki H (2010). A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Alternatives to Laboratory Animals* 38(4): 275-284.

Asturiol D & Worth A (2011). The use of chemical reactivity assays in toxicity prediction. JRC Scientific and Technical Reports. Available from:
<http://publications.jrc.ec.europa.eu/repository/handle/111111111/22180>

Basketter DA, Casati S, Cronin MTD, Diembeck W, Gerberick GF, Hadgraft J, Kasting G, Marty JP, Nikolaidis E, Patlewicz G, Pease C, Roberts DW, Roggen E, Rovida C & van der Sandt J (2007). Skin sensitisation and epidermal disposition: the relevance of epidermal disposition for sensitisation hazard identification and risk assessment. *Alternatives to Laboratory Animals* 35(1): 137-154.

Basketter DA & Kimber I (2010). Contact hypersensitivity: In: McQueen CA (ed) *Comprehensive Toxicology*, 2nd edition. Elsevier, Kidlington. Vol 5: 397-411.

Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B & Landsiedel R (2012). Putting the parts together: combining in vitro methods to test for skin sensitizing potential. *Regulatory Toxicology and Pharmacology* 63: 489-504.

Casati S, Aeby P, Basketter DA, Cavani A, Gennari A, Gerberick GF, Griem P, Hartung T, Kimber I, Lepoittevin JP, Meade BJ, Pallardy M, Rougier N, Rousset F, Rubinstenn G, Sallusto F, Verheyen GR & Zuang V (2005). Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. *Alternatives to Laboratory Animals* 33: 47-62.

Chaudhry Q, Piclin N, Cotterill J, Pintore M, Price NR, Chrétien JR & Roncaglioni A (2010). Global QSAR models of skin sensitizers for regulatory purposes. *Chemistry Central Journal* 4(1):S5.

Dancik Y, Miller MA, Jaworska J & Kasting GB (2013). Design and performance of a spreadsheet-based model for estimating bioavailability of chemicals from dermal exposure. *Advanced Drug Delivery Reviews* 65: 221-236.

EURL ECVAM (2013). EURL ECVAM strategy for replacement of animal testing for skin sensitisation hazard identification and classification. JRC scientific and policy reports. Available from:
<http://publications.jrc.ec.europa.eu/repository/handle/111111111/27708>

European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official Journal of the European Union* L396:1-849.

European Commission (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353: 1-1355.

European Commission (2009a). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309: 1-47.

European Commission (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342: 59-209.

Emter R, Ellis G & Natsch A (2010). Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. *Toxicology and Applied Pharmacology* 245: 281-290.
European Commission, 2006.

European Union (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167: 1-116.

Galbiati V, Lucchi L, Mitjans M, Viviani B, Galli CL, Marinovitch M & Corsini E (2011). Further development of the NCTC 2544 IL-18 assay to identify in vitro contact allergens. *Toxicology in Vitro* 25: 724-732.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG & Lepoittevin JP (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. *Toxicological Sciences* 97(2): 417-427.

Gerberick F, Aleksic M, Basketter D, Casati S, Karlberg AT, Kern P, Kimber I, Lepoittevin JP, Natsch A, Ovigne JM, Rovida C, Sakaguchi H & Schultz T (2008). Chemical reactivity measurement and the predictive identification of skin sensitizers. *Alternatives to Laboratory Animals* 36: 215-242.

Gerberick GF, Troutman JA, Foertsch LM, Vassallo JD, Quijano M, Dobson RL, Goebel C & Lepoittevin JP (2009). Investigation of peptide reactivity of pro-hapten skin sensitizers using a peroxidase-peroxide oxidation system. *Toxicological Sciences* 112(1): 164-174.

Griffiths CEM, Dearman RJ, Cumberbatch M & Kimber I (2005). Cytokines and Langerhans cell mobilisation in mouse and man. *Cytokine* 32: 67-70.

Hirota M, Kouzuki H, Ashikaga T, Sono S, Tsujita K, Sasa H & Aiba S (2013). Artificial neural network analysis of data from multiple in vitro assays for prediction of skin sensitization potency of chemicals. *Toxicology In Vitro* 27(4): 1233-1246.

Jaworska J, Harol A, Kern PS & Gerberick F (2011). Integrating non-animal test information into an adaptive testing strategy – skin sensitization proof of concept phase. *Alternative to Animal Experimentation* 28: 211-225.

Jaworska J, Dancik Y, Kern P, Gerberick F & Natsch A (2013). Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. *Journal of Applied Toxicology* 33:1353-1364.

Johansson H, Albrecht A-S, Borrebaek CAK & Lindstedt M (2013). The GARD assay for assessment of chemical skin sensitizers. *Toxicology in Vitro* 27: 1163-1169.

Jowsey IR, Basketter DA, Westmoreland C & Kimber I (2006). A future approach to measuring relative skin sensitising potency: a proposal. *Journal of Applied Toxicology* 26: 341-350.

Kim, H-J, Barajas B, Wang M & Ne AE (2008). Nrf2 activation by sulforaphane restores the age-related decrease of TH1 immunity. Role of dendritic cells. *Journal of Allergy and Clinical Immunology* 121(5): 1255-1261.

Kimber I & Basketter DA (1992). The murine local lymph node assay: a commentary on collaborative studies and new directions. *Food and Chemical Toxicology* 30(2): 165-169.

Lambrechts N, Vanheel H, Nelissen I, Witters H, van den Heuvel R, Van Tendeloo V, Schoeters G & Hooyberghs J (2010). Assessment of chemical skin-sensitizing potency by an in vitro assay based on human dendritic cells. *Toxicological Sciences* 116(1): 122-129.

Lepoittevin JP (2006). Metabolism versus chemical transformation or pro-versus prehapens? *Contact Dermatitis* 54: 73-74.

Martin SF, Esser PR, Schmucker S, Dietz L, Naisbitt DJ, Park K, Vocanson M, Nicolas JF, Keller M, Pichler WJ, Peiser M, Luch A, Wanner R, Maggi E, Cavani A, Rustemeyer T, Richter A, Thierse HJ & Sallusto F (2010). T-cell recognition of chemicals, protein allergens and drugs: towards the development of in vitro assays. *Cellular and Molecular Life Sciences* 67: 4171-4184.

Maxwell G & MacKay C (2008). Application of a Systems Biology Approach to skin allergy risk assessment. *Alternatives to Laboratory Animals* 36: 521-556.

Maxwell G, MacKay C, Cubberley R, Davies M, Gellatly N, Glavin S, Gouin T, Jacquilleot S, Moore C, Pendlington R, Saib O, Sheffield D, Stark R & Summerfield V (2014). Applying the skin sensitisation adverse outcome pathway (AOP) to quantitative risk assessment. *Toxicology In Vitro* 28(1): 8-12.

McKim JM Jr, Keller DJ 3rd & Gorski JR (2010). A new in vitro method for identifying chemical sensitizers combining peptide binding with ARE/EpRE-mediated gene expression in human skin cells. *Cutaneous and Ocular Toxicology* 29(3): 171-192.

Natsch A, Emter R & Ellis G (2009). Filling the concept with data: Integrating data from different in vitro and in silico assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. *Toxicological Sciences* 107(1): 106-121.

Nukada Y, Miyazawa M, Kazutoshi S, Sakaguchi H & Nishiyama N (2013). Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. *Toxicology in Vitro* 27: 609-618.

Organisation for Economic Cooperation and Development (1992). Guideline for the Testing of Chemicals, 406, Skin sensitisation. Paris, France. Available from: <http://www.oecd.org>

Organisation for Economic Cooperation and Development (2010a). Guideline for the Testing of Chemicals, 429, Skin Sensitization: Local Lymph Node Assay. Paris, France. Available from: <http://www.oecd.org>

Organisation for Economic Cooperation and Development (2010b). Guideline for the Testing of Chemicals, 442A, Skin Sensitization: Local Lymph Node Assay: DA. Paris, France. Available from: <http://www.oecd.org>

Organisation for Economic Cooperation and Development (2010c). Guideline for the Testing of Chemicals, 442B, Skin Sensitization: Local Lymph Node assay: BrdU-ELISA. Paris, France. Available from: <http://www.oecd.org>

Organisation for Economic Cooperation and Development (2012a). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins; Part 1: Scientific Evidence Series on Testing and Assessment No.168 ENV/JM/MONO(2012)10/PART1. Paris, France. Available from: <http://www.oecd.org>

Organisation for Economic Cooperation and Development (2012b). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins; Part 2: Use of the AOP to Develop Chemical Categories and Integrated Assessment and Testing Approaches Series on Testing and Assessment No.168 ENV/JM/MONO(2012)10/PART2. Paris, France. Available from: <http://www.oecd.org>

Patlewicz G, Aptula AO, Roberts DW & Uriarte E (2008). A Minireview of Available Skin Sensitization (Q)SARs/Expert Systems. *QSAR & Combinatorial Science* 27(1): 60-76.

Patlewicz G, Dimitrov SD, Low LK, Kern PS, Dimitrova GD, Comber MIH, Aptula AO, Phillips RD, Niemela J, Madsen C, Wedeby EB, Roberts DW, Bailey PT & Mekenyan OG (2007). TIMES-SS – A promising tool for the assessment of skin sensitization hazard. A characterisation with respect to the OECD Validation Principles for (Q)SARs. *Regulatory Toxicology and Pharmacology* 48: 225–239.

Patlewicz G & Worth A (2008). Review of data sources, QSARs and Integrated Testing Strategies for skin sensitisation. JRC Scientific and Technical Reports. Available from: <http://publications.jrc.ec.europa.eu/repository/handle/111111111/998>

Patlewicz G, Kuseva C, Kesova A, Popova I, Zhechev T, Pavlov T, Roberts DW & Mekenyan O (2014). Towards AOP application - Implementation of an integrated approach to testing and assessment (IATA) into a pipeline tool for skin sensitization. *Regulatory Toxicology and Pharmacology* 69(3): 529-545.

Roberts DW & Williams DL (1982). The derivation of quantitative correlations between skin sensitisation and physio-chemical parameters for alkylating agents, and their application to experimental data for sultones. *Journal of Theoretical Biology* 99(4): 807-825.

Roberts DW, Patlewicz G, Dimitrov SD, Low LK, Aptula AO, Kern PS, Dimitrova GD, Comber MIH, Phillips RD, Niemelä J, Madsen C, Wedeby EB, Bailey PT & Mekenyan OG (2007). TIMES-SS - a mechanistic evaluation of an external validation study using reaction chemistry principles. *Chemical Research in Toxicology* 20(9):1321-1330.

Sakaguchi H, Ashikaga T, Miyazawa M, Yoshida Y, Ito Y, Yoneyama K, Hirota M, Itagaki H, Toyoda H & Suzuki H (2006). Development of an in vitro skin sensitization test using human cell lines; human Cell Line Activation Test (h-CLAT). II. An inter laboratory study of the h-CLAT. *Toxicology in Vitro* 20: 774-84.

Schultz TW, Yarbrough JW & Johnson EL (2005). Structure-activity relationships for reactivity of carbonyl-containing compounds with glutathione. *SAR and QSAR in Environmental Research* 16(4): 313-322.

van der Veen JW, Rorije E, Emter R, Natsch A, van Loveren H & Ezendam J (2014). Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. *Regulatory Toxicology and Pharmacology* 69(3): 371-379.

Vocanson M, Hennino A, Rozieres A, Poyet G & Nicolas JF (2009). Effector and regulatory mechanisms in allergic contact dermatitis. *Allergy* 64: 1699-1714.

WHO (2006). Environmental Health Criteria 235. Dermal absorption. Available from: <http://www.who.int/ipcs/features/2006/ehc235/en/>

Chapter 5. Table of References with Notes

<i>Regulatory requirements across sectors</i>	
<p>European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L396: 1-849.</p>	<p><i>Chemicals</i>: Information requirements depend on the production volume of the chemical concerned. Annex VII: Skin sensitisation testing is mandatory for all substances manufactured or imported in the EU in quantities of one ton or more per year (lowest tonnage level). For new testing of substances the LLNA (OECD TG 429) is the method of first choice.</p>
<p>European Commission (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353: 1-1355.</p>	<p><i>CLP</i>: All substances and mixture placed on the market irrespective of the tonnage level. Currently allows classification of sensitiser into one hazard category. Classification into potency categories is expected to be introduced in future amendments for complete implementation of the Globally Harmonised System of Classification and Labelling.</p>
<p>European Commission (2009a). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309: 1-47.</p>	<p><i>Plant Protection Products (PPPs)</i>: All plant protection products and active substances marketed in the UE independent of tonnage level. Information requirement for skin sensitisation is mandatory for both, active substances contained in PPPs and the final product. The test is not needed if the active substance is classified as a sensitiser or is otherwise known to be a sensitiser (from human data).</p>
<p>European Commission (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342: 59-209.</p>	<p><i>Cosmetics</i>: All cosmetics ingredients independent of tonnage level. Skin sensitisation is one of the most relevant endpoints for the toxicological profile of cosmetic ingredients. In March 2013 a marketing ban for cosmetic products containing ingredients tested on animals entered into force</p>

<p>European Union Regulation (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167: 1-116.</p>	<p><i>Biocidals:</i> All biocidal products and active substances marketed in the UE independent of tonnage level. Information requirement for skin sensitisation mandatory for both active substances contained in biocidal products and the final product. The test is not needed if the active substance is classified as a sensitiser according to Directive 67/548/EEC or is otherwise known to be a sensitiser (from human data).</p>
<p><i>Traditional animal tests – OECD TGs and GD</i></p>	
<p>Organisation for Economic Cooperation and Development (1992). Guideline for the Testing of Chemicals, 406, Skin sensitisation. Paris, France. Available from: http://www.oecd.org.</p>	<p>This Test Guideline describes the Guinea Pig Maximisation Test (GMPT) of Magnusson and Kligman which uses adjuvant to potentiate sensitisation and the non-adjuvant Buehler test. Both methods provide information on the skin sensitisation potential of chemicals in animals.</p>
<p>Organisation for Economic Cooperation and Development (2010a). Guideline for the Testing of Chemicals, 429, Skin Sensitization: Local Lymph Node Assay. Paris, France. Available from: http://www.oecd.org</p>	<p>This Test Guideline describes the Local Lymph Node assay (LLNA). The method is designed for assessing the skin sensitisation potential of chemicals in animals and provides quantitative data suitable for dose-response assessment. The basic principle of this assay is that sensitisers induce a primary proliferation of lymphocytes in the auricular lymph nodes draining the site of chemical application. To measure cell proliferation, this method is based on the use of radioactive labelling.</p>
<p>Organisation for Economic Cooperation and Development (2010b). Guideline for the Testing of Chemicals, 442A, Skin Sensitization: Local Lymph Node Assay: DA. Paris, France. Available from: http://www.oecd.org</p>	<p>This Test Guideline describes a non-radioactive modification to the LLNA method for identifying potential skin sensitizing test substances. The measurement of adenosine triphosphate (ATP) content by bio-luminescence is used as an indicator of lymphocyte proliferation.</p>
<p>Organisation for Economic Cooperation and Development (2010c). Guideline for the Testing of Chemicals, 442B, Skin Sensitization: Local Lymph Node assay: BrdU-ELISA. Paris, France. Available from: http://www.oecd.org</p>	<p>This Test Guideline describes a non-radioactive modification to the LLNA method for identifying potential skin sensitizing test substances. The measurement of 5-bromo-2-deoxyuridine (BrdU) content is used as an indicator of lymphocytes proliferation.</p>
<p><i>Traditional animal tests – Scientific publications</i></p>	
<p>Kimber I & Basketter DA (1992). The murine local lymph node assay: a commentary on collaborative studies and new directions. Food and Chemical Toxicology 30(2): 165-169.</p>	<p>This paper provides a historical background to the development of the murine LLNA. It reports performance of the LLNA in an interlaboratory trial designed to evaluate the utility of the method as an alternative to the guinea-pig methods. In addition the authors make some recommendations regarding the use and interpretation of the LLNA including</p>

	comparisons of the LLNA with guinea-pig methods.
	<i>Mechanistic understanding</i>
Adler S, Basketter D, Creton S, et al. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Archives of Toxicology 85(5): 367-485.	This report provides an analysis made by a panel of experts of the status and prospects of alternative methods in five toxicological areas of concern in view of the full marketing ban entered into force in 2013 for cosmetics products and ingredients tested on animals in Europe. The report includes five chapters dealing with toxicokinetics, skin sensitisation, repeated dose toxicity, carcinogenicity and reproductive toxicity.
Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N & Itagaki H (2010). A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). Alternatives to Laboratory Animals 38(4): 275-284.	This paper describes a comparative evaluation of the h-CLAT <i>in vitro</i> test and the LLNA in order to confirm the predictive value of the h-CLAT for skin sensitisation activity. The study tested one hundred chemicals including non-sensitisers and sensitisers distributed in the four LLNA potency classes: weak, moderate, strong and extreme. The correlation of the h-CLAT results with the LLNA results was 84%. The use of the combination of CD86 and CD54 induction as positive indicator, improved the accuracy of the test. In conclusion, the h-CLAT is expected to be a useful cell-based <i>in vitro</i> method for predicting skin sensitisation potential
Basketter DA & Kimber I (2010). Contact hypersensitivity: In: McQueen CA (ed) Comprehensive Toxicology, 2 nd edition. Elsevier, Kidlington. Vol 5: 397-411.	This article provides a detailed review of the mechanistic aspects of skin sensitisation and contact hypersensitivity.
Griffiths CEM, Dearman RJ, Cumberbatch M & Kimber I (2005). Cytokines and Langerhans cell mobilisation in mouse and man. Cytokine 32: 67-70.	This review focuses on the mobilisation of epidermal Langerhans cells (LC), critical event during the development of cutaneous immune response. The authors highlight that regulation of mobilisation and migration of LC present similarities in man and mouse as shown by experimental studies. They concluded that, with respect to LC biology at least, the mouse provides a very valuable experimental surrogate for the human skin immune system.
Kim H-J, Barajas B, Wang M & Nel AE (2008). Nrf2 activation by sulforaphane restores the age-related decrease of TH1 immunity. Role of dendritic cells. Journal of Allergy and Clinical Immunology 121(5): 1255-1261.	This paper presents the results of an <i>in vivo</i> study in Nrf2-knockout mice demonstrating that the Nrf2 pathway, which regulates the transcriptional activation of more than 200 antioxidant and protective genes, is a relevant pathway for contact allergy.
Lepoittevin JP (2006). Metabolism <i>versus</i> chemical transformation or pro- <i>versus</i> prehapten? Contact Dermatitis 54: 73-74.	In this editorial, the author introduces the term of “prehapten” to define non-reactive sensitising molecules transformed into haptens by simple chemical transformation and without requirement of a specific enzymatic system.

<p>Organisation for Economic Cooperation and Development (2012a). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins; Part 1: Scientific Evidence Series on Testing and Assessment No.168 ENV/JM/MONO(2012)10/PART1. Paris, France. Available from: http://www.oecd.org</p>	<p>This OECD document describes the collective knowledge of the adverse outcome pathway (AOP) for skin sensitisation initiated by covalent binding to proteins. It also provides an overview of the scientific evidence supporting the AOP and of the experimental evidence supporting the key events in the AOP.</p>
<p>Organisation for Economic Cooperation and Development (2012b). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins; Part 2: Use of the AOP to Develop Chemical Categories and Integrated Assessment and Testing Approaches Series on Testing and Assessment No.168 ENV/JM/MONO(2012)10/PART2. Paris, France. Available from: http://www.oecd.org</p>	<p>This document focuses on how to use the AOP in the context of forming chemical categories. In addition, it addresses the efforts to build AOP-based Integrated Approaches to Testing and Assessment for skin sensitisation.</p>
<p>Vocanson M, Hennino A, Rozieres A, Poyet G & Nicolas JF (2009). Effector and regulatory mechanisms in allergic contact dermatitis. <i>Allergy</i> 64: 1699-1714.</p>	<p>This paper provides a detailed review of the mechanisms in allergic contact dermatitis (ACD).</p>
<p><i>Status of non-standard methods</i></p>	
<p>Asturiol D & Worth A (2011). The use of chemical reactivity assays in toxicity prediction. JRC Scientific and Technical Reports. Available from: http://publications.jrc.ec.europa.eu/repository/handle/11111111/22180</p>	<p>In this report the authors illustrate the basis of the <i>in chemico</i> approach to toxicity prediction in the fields of skin sensitisation, aquatic toxicity and hepatotoxicity and review the studies that have developed the concept and its practical application since the 1930s, with special attention paid to studies aimed at the development of Quantitative Structure-Activity Relationship (QSAR) models and read-across approaches. Authors conclude that <i>in chemico</i> approaches provide a promising means of toxicity prediction, especially in the context of integrated testing strategies based on the use of multiple non-animal methods.</p>
<p>Basketter DA, Casati S, Cronin MTD, Diembeck W, Gerberick GF, Hadgraft J, Kasting G, Marty JP, Nikolaidis E, Patlewicz E, Pease C, Roberts DW, Roggen E, Rovida C & van der Standt J (2007). Skin sensitisation and epidermal disposition: the relevance of epidermal disposition for sensitisation hazard</p>	<p>This report describes the outcome of a workshop on skin sensitisation and epidermal disposition aimed to review the state-of-the-art of approaches available to measure the disposition of chemicals in skin compartments and to develop recommendations on how to use such information into non-animal testing strategies for skin sensitisation. The workshop participants acknowledged the importance of epidermal disposition information in the context of Integrated Testing Strategies for both hazard and potency</p>

<p>identification and risk assessment. <i>Alternatives to Laboratory Animals</i> 35(1): 137-154.</p>	<p>prediction. However, they also acknowledged that the requirements for epidermal disposition data may differ depending on whether the need is for hazard identification or for risk assessment.</p>
<p>Böhme A, Thaens D, Paschke A & Schüürmann G (2009). Kinetic glutathione chemoassay to quantify thiol reactivity of organic electrophiles--application to α,β-unsaturated ketones, acrylates, and propiolates. <i>Chemical Research in Toxicology</i> 22(4): 742-750.</p>	<p>This paper describes a new kinetic glutathione (GSH) chemoassay which employs a photometric method to quantify GSH loss enabling an efficient determination of second-order rate constants, k_{GSH} of the reaction between electrophilic substances and GSH. First results obtained with 15 α,β-unsaturated ketones and 11 α,β-unsaturated ketones demonstrate the capability of the kinetic GSH chemoassay to quantify electrophilic reactivity over a large value range and to serve as a respective ITS tool for the analysis and the prediction of the reactive toxicity of electrophiles.</p>
<p>Casati S, Aeby P, Basketter DA, Cavani A, Gennari A, Gerberick GF, Griem P, Hartung T, Kimber I, Lepoittevin JP, Meade BJ, Pallardy M, Rougier N, Rousset F, Rubinstenn G, Sallusto F, Verheyen GR & Zuang V (2005). Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. <i>Alternatives to Laboratory Animals</i> 33: 47-62.</p>	<p>This report presents the outcome of a workshop aimed to review the state-of-the-art of the use of cultured dendritic cells for the identification of skin sensitisation hazard and to develop strategies for the eventual replacement of <i>in vivo</i> testing. The authors concluded that opportunities are available for exploring the potential of cultured DCs and cell lines of hematopoietic origin in order to provide the basis for <i>in vitro</i> skin sensitisation testing. General observations and recommendations are highlighted for further development, subsequent evaluation and validation and requirements for further research. Progress has been made in the area since the publication of this workshop report, standardised test methods based on the use of DC-like cell lines are now available.</p>
<p>Chaudhry Q, Piclin N, Cotterill J, Pintore M, Price NR, Chrétien JR & Roncaglioni A (2010). Global QSAR models of skin sensitisers for regulatory purposes. <i>Chemistry Central Journal</i> 4(1):S5</p>	<p>This paper reports the development of two global QSAR models. The construction of the models was based on a common dataset of 209 heterogeneous compounds using two computational techniques, particularly suited for data presented as classes of responses.</p>
<p>Dancik Y, Miller MA, Jaworska J & Kasting GB (2013). Design and performance of a spreadsheet-based model for estimating bioavailability of chemicals from dermal exposure. <i>Advanced Drug Delivery Reviews</i> 65: 221-236.</p>	<p>This paper presents an <i>in silico</i> model of chemical penetration through the stratum corneum, viable epidermis and dermis formulated in term of an Excel™ spread sheet. Model predictions are compared with representative <i>in vitro</i> skin permeation data obtained from the literature using as summary parameters total absorption (Q_{abs}), maximum flux (J_{max}) and skin permeability coefficient (k_p). The results of this evaluation demonstrate the current state-of-the-art in prediction of transient skin absorption and highlight areas in which further elaborations are needed to obtain satisfactory predictions.</p>
<p>EURL ECVAM (2013). EURL ECVAM strategy for replacement of animal testing for skin sensitisation hazard identification and classification. JRC scientific</p>	<p>This document outlines the EURL ECVAM strategy for the skin sensitisation area. In order to achieve the highest impact on the 3Rs ECVAM will focus its efforts in the next five years on the development of non-animal testing strategies for hazard identification</p>

and policy reports. Available from: http://publications.jrc.ec.europa.eu/repository/handle/111111111/27708	and sub-categorisation according to sub-categories 1A and 1B of the Globally Harmonised System for Classification and Labelling of Chemicals (GHS). In addition EURL ECVAM will support the OECD in the development of generic guidance on how to document and evaluate Integrated Approaches to Testing and Assessment (IATA).
Emter R, Ellis G & Natsch A (2010). Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers <i>in vitro</i> . <i>Toxicology and Applied Pharmacology</i> 245: 281-290.	This study reports the assessment of the performance of a reporter cell line derived from keratinocytes stably transfected with a luciferase reporter gene under the control of a single copy of antioxidant response element (ARE)-element of the human <i>AKR1C2</i> gene for assessing the skin sensitisation potential of substances.
Galbiati V, Lucchi L, Mitjans M, Viviani B, Galli CL, Marinovitch M & Corsini E (2011). Further development of the NCTC 2544 IL-18 assay to identify <i>in vitro</i> contact allergens. <i>Toxicology in Vitro</i> 25: 724-732.	This paper describes the work undertaken to optimise a test method based on the quantification of IL-18 production in the human keratinocyte cell line NTC 2544 to discriminate contact sensitizer from irritants and low molecular weight respiratory allergens.
Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG & Lepoittevin JP (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. <i>Toxicological Sciences</i> 97(2): 417-427.	In this study, the authors tested the ability of 82 chemicals (including non-allergens and allergens of different potencies) to react with reduced glutathione or with two synthetic peptides, containing either cysteine or lysine, in order to determine whether and to what extent peptide reactivity correlates with skin sensitisation potential. The results showed that measurement of peptide reactivity has considerable potential utility as a screening approach for skin sensitisation testing, and thereby for reducing reliance on animal-based test methods.
Gerberick F, Aleksic M, Basketter D, Casati S, Karlberg AT, Kern P, Kimber I, Lepoittevin JP, Natsch A, Ovigne JM, Rovida C, Sakaguchi H & Schultz T (2008). Chemical reactivity measurement and the predictive identification of skin sensitizers. <i>Alternatives to Laboratory Animals</i> 36: 215-242.	This report presents the outcome of a workshop aimed to review the state-of-the-art of methods for the identification of skin sensitizers based on measurement of chemical reactivity including their limitations. Furthermore, consideration was given as to how such methods could contribute to integrated testing strategies for the replacement of <i>in vivo</i> testing. A number of recommendations were made in order to promote the progress of relevant and reliable methods towards prevalidation and validation.
Gerberick GF, Troutman JA, Foertsch LM, Vassallo JD, Quijano M, Dobson RLM, Goebel C & Lepoittevin JP (2009). Investigation of peptide reactivity of pro-hapten skin sensitizers using a peroxidase-peroxide oxidation system. <i>Toxicological Sciences</i> 112(1): 164-174.	This paper describes the use of an enzyme-mediated activation step (horseradish peroxidase and hydrogen peroxide) into an <i>in chemico</i> skin sensitisation assay (peptide reactivity) for assessing the skin sensitisation potential of pro-haptens.
Johansson H, Albrecht A-S, Borrebaek CAK &	This paper describes an assay for skin sensitisation hazard assessment which is based on a

<p>Lindstedt M (2013). The Gard assay for assessment of chemical skin sensitizers. <i>Toxicology in Vitro</i> 27: 1163-1169.</p>	<p>genomic readout. A signature of 200 genes in the myeloid cell line MUTZ-3 is used to discriminate between sensitizers and non sensitizers. The genes are reported to participate in signalling pathways that are involved with recognition of foreign substances. A number of these pathways, such as nuclear factor-erythroid-related factor 2 (NRF2) mediated oxidative response, aryl hydrocarbon receptor (AHR) signalling and Toll-like receptor (TLR) signalling, are known to lead to transcription of cytoprotective enzymes and DC maturation as a response to xenobiotic challenges.</p>
<p>Lambrechts N, Vanheel H, Nelissen I, Witters H, van den Heuvel R, Van Tendeloo V, Schoeters G & Hooyberghs J (2010). Assessment of chemical skin-sensitizing potency by an <i>in vitro</i> assay based on human dendritic cells. <i>Toxicological Sciences</i> 116(1): 122-129.</p>	<p>In the study described in this paper, authors tried to predict the potency (LLNA EC3 values) of 15 skin sensitizers using <i>in vitro</i>-generated gene expression data from CD34-DC-based assay associated to the concentration of the compound that causes 20% (IC20) cell damage in CD34-DC. Applying a linear regression with both IC20 and expression changes of CREM and CCR2, a high linear correlation was established between the <i>in vitro</i> model and the LLNA values.</p>
<p>Martin SF, Esser PR, Schmucker S, Dietz L, Naisbitt DJ, Park K, Vocanson M, Nicolas JF, Keller M, Pichler WJ, Peiser M, Luch A, Wanner R, Maggi E, Cavani A, Rustemeyer T, Richter A, Thierse HJ & Sallusto F (2010). T-cell recognition of chemicals, protein allergens and drugs: towards the development of <i>in vitro</i> assays. <i>Cellular and Molecular Life Sciences</i> 67: 4171-4184.</p>	<p>This review reports the outcome of an expert meeting aimed to review the status of development of T-cell-based <i>in vitro</i> assays as tools in immunotoxicology to identify hazardous chemicals and drugs.</p>
<p>Maxwell G & MacKay C (2008). Application of a Systems Biology Approach to skin allergy risk assessment. <i>Alternatives to Laboratory Animals</i> 36: 521-556.</p>	<p>This paper describes an <i>in silico</i> model of the induction of skin sensitisation. This model was developed in order to characterise and quantify the contribution of each pathway to the overall biological process. Using a “systems biology” approach, a computer-based mathematical model was constructed. In this one, biological mechanisms underlying the induction phase of skin sensitisation are represented by non-linear ordinary differential equation and defined by using from over 500 published papers. By using this model, authors identified knowledge gaps for future investigation and key factors with major influence on the induction of skin sensitisation.</p>
<p>McKim JM Jr, Keller DJ 3rd & Gorski JR (2010). A new <i>in vitro</i> method for identifying chemical sensitizers combining peptide binding with ARE/EpRE-mediated gene expression in human skin cells. <i>Cutaneous and</i></p>	<p>This study reports the development of a new <i>in vitro</i> screening assay using a human skin cell line (HaCat), chemical reactivity (measurement of glutathione depletion in a cell free matrix) and gene expression profiling for three signalling pathways (Keap1/Nrf2/ARE/EpRE, ARNT/AhR/XRE and Nrf1/MTF/MRE) known to be activated by</p>

Ocular Toxicology 29(3): 171-192.	sensitizing agents. The system was developed with 39 chemicals and 3 cationic metals and challenged with 58 additional compounds. With an accuracy of approximately 84%, a sensitivity of 81% and a specificity of 92%, authors concluded the method described demonstrates a valuable <i>in vitro</i> method to identify chemicals and metals that induce skin sensitisation.
Patlewicz G, Dimitrov SD, Low LK, Kern PS, Dimitrova GD, Comber MIH, Aptula AO, Phillips RD, Niemela J, Madsen C, Wedebye EB, Roberts DW, Bailey PT & Mekenyan OG (2007). TIMES-SS – A promising tool for the assessment of skin sensitization hazard. A characterisation with respect to the OECD Validation Principles for (Q)SARs. Regulatory Toxicology and Pharmacology 48: 225–239.	This study reports the results of an external evaluation of the TIMES-SS, a hybrid expert system encoding structure-toxicity and structure-skin metabolism relationships through a number of transformations. As part of this exercise data were generated for 40 chemicals in the LLNA and then compared with predictions made by TIMES-SS with an overall concordance of 83%. In addition the extent of adherence of the model to the five OECD principles for (Q)SAR validation was evaluated. The authors concluded that the model is a promising tool to aid in the evaluation of skin sensitisation hazard under legislative programs such as REACH.
Patlewicz G, Aptula AO, Roberts DW & Uriarte E (2008). A Minireview of Available Skin Sensitization (Q)SARs/Expert Systems. QSAR & Combinatorial Science 27(1): 60-76.	This minireview describes the state-of-the-art of available (Q)SARs/expert systems for skin sensitisation and evaluates their utility for potential regulatory use.
Patlewicz G & Worth A (2008). Review of data sources, QSARs and Integrated Testing Strategies for skin sensitisation. JRC Scientific and Technical Reports. Available from: http://publications.jrc.ec.europa.eu/repository/handle/11111111/998	This review reports information of skin sensitisation data and computational tools for the estimation of skin sensitisation potential. It presents an overview of what constitutes an Integrated Testing Strategy. Summaries of general models, models for specific chemical classes and mechanisms of action and expert system are provided.
Roberts DW & Williams DL (1982). The derivation of quantitative correlations between skin sensitisation and physio-chemical parameters for alkylating agents, and their application to experimental data for sultones. Journal of Theoretical Biology 99(4): 807-825.	This paper describes a new mathematical model based on quantitative correlation between skin sensitisation and physio-chemical parameters for alkylating agents: the relative alkylating index RAI is intended to provide a relative measure of the extent of <i>in vivo</i> alkylation and is function of chemical reactivity, lipophilicity and dose administrated.
Roberts DW, Patlewicz G, Dimitrov SD, Low LK, Aptula AO, Kern PS, Dimitrova GD, Comber MIH, Phillips RD, Niemelä J, Madsen C, Wedebye EB, Bailey PT & Mekenyan OG (2007). TIMES-SS--a mechanistic evaluation of an external validation study	This paper describes the evaluation of LLNA results, generated in a study designed to perform an external evaluation of the Times Metabolism Simulator platform (TIMES-SS) for predicting skin sensitisation, with respect to reaction chemistry principles for sensitization. Testing on additional four chemicals was carried out to explore some of the specific reaction chemistry findings in more detail. Improvements for TIMES-SS, where

using reaction chemistry principles. <i>Chemical Research in Toxicology</i> 20(9):1321-1330.	appropriate, were put forward together with proposals for further research work.
Schultz TW, Yarbrough JW & Johnson EL (2005). Structure-activity relationships for reactivity of carbonyl-containing compounds with glutathione. <i>SAR and QSAR in Environmental Research</i> 16(4): 313-322.	This paper describes the development of a spectrophotometric assay for determining the reactive to glutathione (GSH) and its use to determine GSH reactivity of 21 aliphatic derivatives of esters, ketones and aldehydes. To quantify the reactivity, the parameter chosen was the RC50 (concentration giving 50% depletion of GSH) calculated from concentration-response curves and which are algebraically related to the rate constant and independent of the experimental methodology. Results were then used to evaluate a series of structure-activity relationship.
<i>Integrated Testing Strategies</i>	
Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B & Landsiedel R (2012). Putting the parts together: combining <i>in vitro</i> methods to test for skin sensitizing potential. <i>Regulatory Toxicology and Pharmacology</i> 63: 489-504.	In this paper the authors report the performance of different <i>in vitro</i> (h-CLAT, LuSens assay, Keratinosens and mMUSST), <i>in chemico</i> (DPRA) and <i>in silico</i> (OECD QSAR toolbox) methods in predicting the skin sensitisation potential of a set of 54 test substances compared to human or LLNA data. The predictive performance of test methods' combinations has also been explored and an <i>in vitro</i> testing scheme and a prediction model based on information on peptide reactivity, activation of Keap-1/Nrf2 signalling pathway and dendritic cell activation is proposed. For the set of 54 chemicals evaluated in the study the model offers an overall accuracy of 94% (sensitivity 93% and specificity 95%) compared to human data and an overall accuracy of 83% (sensitivity 81% and specificity 88%) when compared to LLNA data. The authors conclude that the high accuracy of the proposed prediction model points to a possible future replacement of animal testing while maintaining the same degree of accuracy and prediction for human skin sensitisers. Future activities should focus on better defining the applicability domain of the test battery and explore its potential to deliver potency information.
Goebel C, Aeby P, Ade N, et al. (2012). Guiding principles for the implementation of non-animal safety assessment approaches for cosmetics: Skin sensitisation, <i>Regulatory Toxicology & Pharmacology</i> 63: 40-52.	This paper reports the outcome of a workshop aimed to analyse to what extent skin sensitisation safety assessments for cosmetic ingredients can be made in absence of animal data. It was evaluated how and when non-animal test methods, predictions based on physico-chemical properties, threshold concepts and weight-of-evidence based hazard characterisation could be used to enable safety decisions. The goal was to propose guiding principles for the application and further development of non-animal safety assessment strategies.
Hirota M, Kouzuki H, Ashikaga T, Sono S, Tsujita K,	This paper describes the development the "iSENSver.1" model to predict LLNA

<p>Sasa H & Aiba S (2013). Artificial neural network analysis of data from multiple in vitro assays for prediction of skin sensitization potency of chemicals. <i>Toxicology In Vitro</i> 27(4): 1233-1246.</p>	<p>thresholds (i.e. EC3 values for LLNA-positive chemicals and maximum concentrations for LLNA-negative chemicals) of chemicals from in vitro test data, using the learning function of a non-linear artificial neural network (ANN). MAC value (SH test). The minimum EC150, EC200 (h-CLAT) and CV75 (cytotoxicity) were selected as descriptors for the input layer of the ANN-system. The ANN-predicted values were well correlated with reported LLNA thresholds. The authors concluded that the use of ANN appears be promising for predicting skin sensitisation potential of chemicals from in vitro test data.</p>
<p>Jaworska J, Harol A, Kern PS & Gerberick F (2011). Integrating non-animal test information into an adaptive testing strategy – skin sensitization proof of concept phase. <i>Alternative to Animal Experimentation</i> 28: 211-225.</p>	<p>The paper describes the development of a Bayesian Network Integrated Testing Strategy (BN ITS) for skin sensitization hazard assessment and with the specific goal of estimating potency in the mouse LLNA. The BN ITS combines biological knowledge with heterogeneous experimental <i>in silico</i>, <i>in chemico</i>, and <i>in vitro</i> data related to skin penetration, peptide reactivity, and dendritic cell activation. Prediction results are expressed in terms of probabilistic hypothesis about the potency of chemical as predicted by the LLNA. The BN ITS is proposed as a reduction strategy since chemicals with clear potency (i.e. with a high probability associated to the prediction) can be discriminated from chemicals for which more evidence is needed. In addition the BN ITS can be considered a testing guiding tool since it allows the possibility of evaluating the impact of generating additional data on the target information uncertainty reduction before testing is commenced.</p>
<p>Jaworska J, Dancik Y, Kern P, Gerberick F & Natsch A (2013). Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. <i>Journal of Applied Toxicology</i>. <i>Journal of Applied Toxicology</i> 33:1353-1364.</p>	<p>This paper presents an updated Integrated Testing Strategy in the form of a Bayesian Network (BN ITS) developed to assess skin sensitisation potency expressed as LLNA potency classes. The parameters of the updated BN ITS were calculated from an extend data set of 124 chemicals. The improved BN ITS predicted correctly 95% and 86% of chemicals in a test set (n=21) for hazard and LLNA potency classes, respectively. Moreover, the BN ITS model can develop a hypothesis using subsets of data as small as one data point and can be queried on the value of adding additional tests before testing is commenced.</p>
<p>Jowsey IR, Basketter DA, Westmoreland C & Kimber I (2006). A future approach to measuring relative skin sensitising potency: a proposal. <i>Journal of Applied Toxicology</i> 26: 341-350.</p>	<p>In this paper, the authors propose a theoretical framework for the identification of skin sensitising chemicals and for relative potency prediction. The approach is based on the integration of five independent sources of information: structural alerts from SAR and/or expert systems, estimations of epidermal bioavailability, protein reactivity, impact on the phenotype and/or function of DC or DC-like cells and ability to provoke T lymphocyte responses. Specific scoring criteria were</p>

	applied to facilitate both hazard identification and the deviation of a single measure of relative potency.
Maxwell G, MacKay C, Cubberley R, Davies M, Gellatly N, Glavin S, Gouin T, Jacquoilleot S, Moore C, Pendlington R, Saib O, Sheffield D, Stark R & Summerfield V (2014). Applying the skin sensitisation adverse outcome pathway (AOP) to quantitative risk assessment. <i>Toxicology In Vitro</i> 28(1): 8-12.	This article describes the development of two mathematical models ('total haptened protein' model and 'CD8+ T cell response' model) that will be linked in order to provide predictions of the human CD8+ T cell response for a defined skin exposure to a sensitising chemical. Using these approaches, authors intended to quantify the correlation between the dose of sensitiser applied to the skin and the extent of the hapten-specific T cell response that would result. Moreover, using clinical dataset as benchmark points (e.g. human diagnostic patch test data) the authors claimed that this approach could represent a new paradigm for mechanistic toxicology.
Mekenyan O, Patlewicz P, Dimitrova G, Kuseva C, Todorov M, Stoeva S, Kotov S & Donner EM (2010). Use of Genotoxicity Information in the Development of Integrated Testing Strategies (ITS) for Skin Sensitization. <i>Chemical Research in Toxicology</i> 23: 1519-1540.	In this study, the Tissue Metabolism Simulator (TIMES) hybrid expert system was used to compare underlying mechanisms of mutagenicity and skin sensitisation. The authors concluded that mutagenicity information, when used as part of an integrated testing strategy, can provide useful insights on skin sensitisation potential.
Natsch A, Emter R & Ellis G (2009). Filling the concept with data: Integrating data from different <i>in vitro</i> and <i>in silico</i> assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. <i>Toxicological Sciences</i> 107(1): 106-121.	The aim of this paper is to fill for the first time and to use the widely cited "battery approach" proposal of Jowsey et al. 2006 with both novel and literature data compilation on 116 chemicals. The results were scaled into five classes to give an <i>in vitro</i> score and compared to the LLNA results also scaled. Then different ways of data integration were explored to rate the hazard of chemicals and to assess their potency. The optimized model developed gave an overall accuracy for predicting sensitizers of 87.9%, a sensitivity of 85.7%, a specificity of 93.8% and showed a linear correlation between LLNA score and the <i>in vitro</i> score. However, as suggested by the relatively high variation in the <i>in vitro</i> score between chemicals belonging to the same sensitization potency class, the correlation need to be improved.
Nukada Y, Miyazawa M, Kazutoshi S, Sakaguchi H & Nishiyama N (2013). Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. <i>Toxicology in Vitro</i> 27: 609-618.	The authors used a dataset of 101 chemicals with LLNA, h-CLAT, DPRA and <i>in silico</i> prediction system for the development of a non-animal test battery to predict the skin sensitizing potential and potency of chemicals. The results of these studies were scored and compared to LLNA data. The sum of individual scores provided the accuracy of 85% and 71% for the potential and potency prediction. The second tiered system of h-CLAT and DPRA gave 86% and 73%, respectively. It showed a higher sensitivity (96%) compared with h-CLAT alone. These data support that h-CLAT can be a part of a test

	battery. Moreover, they demonstrate the utility of a tiered system when h-CLAT and DPRA are the first screening methods for skin sensitisation.
Patlewicz P, Mekenyan O, Dimitrova G, Kuseva C, Todorov M, Kotov S, Stoeva S & Donner EM (2010). Can mutagenicity information be useful in an Integrated Testing Strategy (ITS) for skin sensitization? SAR and QSAR in Environmental Research 21(7-8): 619-656.	In this study, the authors tested the validity of a relationship between <i>in vitro</i> mutagenicity information and assessment of skin sensitisation potential using a dataset reported by Wolfreys and Basketter (Cutaneous and Ocular Toxicology 23 (2004):197-205) and the Tissue Metabolism Simulator (TIMES). They concluded that mutagenicity information can play a significant role in evaluating sensitisation potential as part of an integrated testing strategy. Careful attention needs to be made to ensure that any information is interpreted in the appropriate context.
Patlewicz G, Kuseva C, Kesova A, Popova I, Zhechev T, Pavlov T, Roberts DW & Mekenyan O (2014). Towards AOP application - Implementation of an integrated approach to testing and assessment (IATA) into a pipeline tool for skin sensitization. Regulatory Toxicology and Pharmacology 69(3): 529-545.	This article describes the development of an Integrated Approach to Testing and Assessment (IATA) for skin sensitisation, implemented into a pipeline tool using OASIS technology. It focuses on existing information including non-testing approaches such as QSAR and read-across. Assessed with a set of 100 substances, the IATA based on <i>in silico</i> and <i>in chemico</i> profiling information showed a preliminary accuracy of 73.5%. Including information from other relevant endpoint coupled with a reaction chemistry mechanistic understanding, accuracy increased (87.6%). Authors concluded that this pipeline platform could be useful in the assessment of skin sensitisation potential.
van der Veen JW, Rorije E, Emter R, Natsch A, van Loveren H & Ezenam J (2014). Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. Regulatory Toxicology and Pharmacology 69(3): 371-379.	This article focuses on the comparison between a tiered strategy (including QSAR, peptide binding, KeratinoSens/Gene Signature and h-CLAT) and a majority voting approach (peptide binding, gene signature or KeratinoSens and h-CLAT). The tiered strategy was able to correctly identify all the 41 chemicals tested. Comparing both methods, the tiered strategy required less experiments but it is more complex in term of different alternative methods required. Costs are similar for both strategies and both provide a mechanistic basis for skin sensitisation testing.

6. Acute systemic toxicity

Pilar Prieto & Andrew Worth

6.1 The traditional animal test and its regulatory use

The term acute systemic toxicity comprises the general adverse effects that occur after a single or multiple exposure of an animal to a substance within 24 hours and during an observation period of at least 14 days. The substance may be administered orally, by inhalation or dermally. Currently, acute oral systemic toxicity is assessed in rats according to three refinement and reduction alternative methods (modifications of the classical LD₅₀ test) described in OECD test guidelines (OECD, 2001a, b, c) and EU Test Methods (EC, 2008, EU, 2014). The endpoint measured in these standard assays is animal morbidity or death (TG423, EU Test Method B.1 tris, TG425) or evident toxicity (TG420, EU Test Method B.1 bis). Clinical signs and conditions associated with pain, suffering, and impending death, are described in detail in the OECD Guidance Document 19 (OECD, 2000). Acute dermal systemic toxicity is assessed according to the classical dermal LD₅₀ study (OECD, 1987) and acute inhalation toxicity according to the revised version of the classical LC₅₀ study (OECD, 2009a, EU Test Method B.2) and the new Acute Toxic Class test guideline (OECD, 2009b, EU Test Method B.52) that allows to reduce considerably the number of animals used (6 to 9 animals instead of 40-80 with the revised TG 403). One of the main purposes of conducting the test is to categorise substances according to their potential hazards and the dose required to cause toxicity (i.e. classification and labelling). The use of lethality as an endpoint has long been criticised on animal welfare grounds, and the utility of the data generated by acute toxicity tests with regard to their ultimate purpose, namely to predict the human hazard potential of substances, has also been questioned (Creton et al., 2010; Seidle et al., 2010).

Regulatory requirements for the test vary depending on the type of chemical under regulation and the region (Seidle et al., 2010). For cosmetics ingredients and products, acute toxicity testing on animals is prohibited in the EU (EC, 2009a), while the test is required for industrial chemicals (EC, 2006), biocides (EU, 2012) and plant protection products (EC, 2009b; EC, 2013). The pharmaceutical sector has discontinued the use of stand-alone acute toxicity studies (ICH, 2009), since it has been shown that they are not useful in supporting the first clinical trial in humans (Robinson et al., 2008) and do not provide additional value in supporting the management of overdose (Chapman et al., 2010). In many circumstances, the relevant information can be obtained from other studies that are already part of drug development such as short-term or dose range finding studies (Robinson et al., 2008, ICH, 2009).

6.2 Mechanistic understanding of the endpoint

What is our current knowledge about human acute toxicity pathways? Is death induced by acute toxic substances the result of specific mechanisms of action involving specific molecular targets in certain tissues, or is it due to general cytotoxicity? What about sublethal effects?

A general conclusion from the Multicenter Evaluation of *In Vitro* Cytotoxicity (MEIC) programme was that the majority of chemicals are acutely toxic to humans by basal cytotoxicity, that is, by interfering with general cell functions common to all cells (Ekwall, 1999). Mechanisms common to many cell types leading to organ failure include, for example, disruption of membrane structure or function, inhibition of mitochondrial function, disturbance of protein turnover, and disruption of metabolism and energy production.

In an ICCVAM/ECVAM/JaCVAM workshop on acute chemical safety testing (NIH, 2009) approaches that could help to identify key toxicity pathways for acute oral toxicity were discussed. It was agreed that the mechanistic information could be used to develop more predictive *in vitro* test methods and to identify earlier and more humane endpoints during *in vivo* testing. However, a major gap at present is our understanding of all *in vivo* mechanisms of action of chemicals. *In vivo* omics evaluation was considered a relevant approach to assess mechanism of action and to identify biomarkers (Boekelheide and Schuppe-Koistinen, 2012). Recently, Liu et al (2013) have integrated the outcome of placental mRNA expression analysis and serological proteome profiling to develop biomarkers of preeclampsia, a pregnancy-related vascular disorder that is the leading cause of maternal morbidity and mortality. Moreover, the collection of as much data as possible from the animal study could help to understand modes of action. Information from diagnosis and treatment of human poisoning could also contribute to improve our understanding and rationale for developing *in vitro* methods for acute systemic toxicity testing. In this regard, the following key events were identified as important to better understand the toxicity of chemicals and to better treat human poisonings: general cellular function, neuronal transmission (central and peripheral), sodium/potassium ATPase function, xenobiotic and aerobic metabolism, cardiac conduction, oxidative stress, receptor activity and immune function (NIH, 2009; see section 6.0 of the report).

While the identification of cellular pathways is relevant in understanding the key initial interactions (e.g. oxidative stress, loss of membrane function, specific interaction with a key receptor), this will not be enough to decipher the subsequent steps involved in acute systemic toxicity pathways. Test methods should allow measurements of integrated cellular responses (e.g. glutathione concentration, mitochondrial function, cytotoxicity, apoptosis, proliferation) or interaction with cellular targets (e.g. over-expression of transporters in cell lines, uptake rates into the cells). The assessment of the interactions between tissues is another challenge when using *in vitro* methods (e.g. an inflammatory response following toxicity in one organ or remote tissue target). Some recommendations of the workshop include the development of MoA-based test methods, the use of human cell-based systems, to study the interaction among the affected cellular pathways assessed in the developed human cell-models, and to develop tools for determining toxicokinetics both *in vitro* and *in vivo*. Prediction of *in vivo* acute toxicity from the activation of specific toxicity pathways requires the development of computational systems biology approaches, in which a pathway-based point of departure (e.g. significant upregulation of mRNA expression of a given gene or gene set) is translated into a corresponding apical point of departure (e.g. oral NOAEL) by physiologically-based biokinetic (PBBK) modelling. In applying this approach, it is first necessary to identify a toxicologically

relevant point of departure at the pathway level, for example, a key event (upstream of the apical effect) in an Adverse Outcome Pathway.

Acute toxicity after oral, dermal or inhalation exposure requires that the substance becomes bioavailable at the target site, which means that biokinetic factors are also key determinants of toxicity (Adler et al., 2011; Coecke et al., 2013). In addition, if the damage involves interference with homeostatic signalling mechanisms, non-exposed tissues and vital organs can also be affected (Gennari et al., 2004). For example, respiratory depression leading to death may be due to depression of the central nervous system rather than a direct effect on the respiratory system.

6.3 Status of non-standard methods and Integrated Testing Strategies

Cytotoxicity assays have been developed and evaluated against *in vivo* oral LD₅₀ data (correlative approaches) and the results of several international projects are published (Ekwall, 1999, NIH, 2006, Prieto et al., 2013a; Kinsner-Ovaskainen et al., 2013). The overall accuracy of the validated neutral red uptake (NRU) basal cytotoxicity test methods when predicting the acute oral toxicity classification categories of the United Nations Globally Harmonised System turned out to be low (~30%). Nevertheless, these methods have been proposed for use in weight-of-evidence approaches to determine the starting dose for acute oral toxicity testing *in vivo* (NIH, 2006, OECD, 2010). Schrage et al (2011) confirmed the low overall concordance of 35% and also reported that the prediction of the starting dose for the subsequent *in vivo* test was useful with regard to potential reduction in animal usage for 59% of the 203 substances tested.

A range of Quantitative Structure-Activity Relationships (QSAR) models have been developed to predict acute systemic toxicity (LD₅₀ values or classifications) via various routes, as reviewed by Lapenna et al (2010). Such models are generally statistically-based models, supported by little or no mechanistic rationale, but they may nevertheless be predictive for specific chemical classes. In a recent JRC study (Norlén et al, 2012), four QSAR models and the 3T3 NRU *in vitro* prediction model were applied to a dataset of 180 chemicals. The results showed that the QSAR models have a predictive capacity (correlation with LD₅₀ of 49-84%, depending on the model) equivalent to or greater than the 3T3 NRU *in vitro* test (correlation with LD₅₀ of ~50%). Similar predictive performances have been obtained in the context of the Antares project (<http://www.antares-life.eu/>; E. Benfenati, personal communication), using a much larger dataset of 7420 compounds. In the JRC study, the five methods were also applied in combination, but this did not yield a significantly better predictive performance. It did however enable the classification of chemicals to be optimised for overall accuracy, sensitivity or specificity, according to the end-user's requirements.

Assuming that most industrial chemicals are not likely to be acutely toxic (Bulgheroni et al., 2009), the capacity of the 3T3 NRU cytotoxicity assay to specifically identify non-classified chemicals on the basis of the 2000 mg/kg b.w. threshold introduced by the EU Classification, Labelling and Packaging (CLP) Regulation (EU CLP 2008) has been evaluated by EURL ECVAM (Prieto et al, 2013b). The results have shown that the test method has a high sensitivity (about 95%) and hence a low false negative rate (less than

5%), which means that substances found to be negative (non-classified) in this test would most likely not require classification for acute oral toxicity under the CLP Regulation. Conversely, the low specificity of the assay (about 42%) means that the test overpredicts many negatives as positives (false positive rate of 58%), which means that the test cannot be used on its own for the identification of substances requiring classification. When the 2000 mg/kg threshold was applied to the data set from the NICEATM/ECVAM validation study (NIH, 2006) and from Schrage et al., (2011), which cover a wide range of chemical uses, a similar high sensitivity and low negative rate was obtained. In the light of these results, it is questionable whether many acutely toxic chemicals really act via a specific mechanism without also showing any cytotoxicity.

Moreover, Bulgheroni and colleagues (2009) evaluated the possibility to identify non-toxic substances (those with an $LD_{50} > 2000$ mg/kg) from 28-days repeated dose studies and set a NOAEL level threshold equal of higher than 200 mg/kg b.w. which allowed to correctly identify 63% (913/1436) of the non-toxic substances. Using this threshold, less than 1% of harmful compounds were misclassified as non-toxic. Due to the 37% of non-toxic substances predicted as toxic (false positives) the ability to predict toxic substances with this approach is difficult.

The EU FP6 ACuteTox project aimed to develop a non-animal testing strategy for predicting human acute oral toxicity by evaluating and combining cytotoxicity assays, organ-specific toxicity assays, and biokinetic/metabolism methods (Kinsner-Ovaskainen et al., 2013). The outcome of this research project reinforced previous results obtained with the 3T3 NRU assay, supporting its use within a tiered testing strategy to identify non-classified substances. The results showed that complementing the 3T3 NRU cytotoxicity assay with additional target organ specific *in vitro* assays did not significantly improve the classification of compounds in acute oral CLP toxicity categories 1-4. However, target organ specific *in vitro* assays were considered to provide added value by alerting for specific toxicity (e.g. neurotoxicity) and further reducing the number of under-predictions (Prieto et al., 2013a; Zurich et al., 2013). In general, the ACuteTox strategies had a tendency to over-predict toxicity, which was explained by their failure to capture *in vivo* biokinetics (i.e. restricted access to the target tissue, quick elimination or detoxification through metabolism). The importance of incorporating suitable kinetic parameters in testing strategies for acute oral toxicity is widely acknowledged (Adler et al., 2011; Coecke et al., 2012; Blaauboer et al., 2012).

One of the efforts of the European Partnership for Alternative Approaches to Animal Testing (EPAA) in the area of acute systemic toxicity is to identify opportunities to avoid redundant testing (EPAA, 2010; 2012). The value of acute toxicity testing by more than one route has been evaluated in several publications (Indans et al., 1998; Thomas and Dewhurst 2007; Seidle et al., 2010; Creton et al., 2010; Seidle et al., 2011; Andrew and Wright-Williams, 2011; Moore et al., 2013). These analyses have shown that testing for acute dermal toxicity is redundant for substances not classified for acute oral toxicity, which casts doubt on the need for acute dermal toxicity as a default information requirement in safety assessment and CLP. The data requirements under the recently adopted Biocidal Products Regulation (European Union, 2012) state that testing via other

routes is only necessary when specified criteria are met. Moreover, the new data requirements for plant protection product active substances allow to waive the dermal route if scientifically justified e.g. where oral $LD_{50} > 2000$ mg/kg (European Commission, 2013). A document containing the technical progress made in some areas has been recently prepared by an EPAA technical expert group with the ultimate goal to invite the European Commission to consider similar improvements in other pieces of EU legislation (e.g. REACH). The Competent Authorities for REACH and CLP (CARACAL) have discussed this proposal to modify REACH standard information requirements for acute toxicity and, at the July 2014 meeting, agreed to amend REACH Annex VIII (point 8.5.3) so that substances that have not shown oral acute toxicity up to a limit dose of 2000mg/kg body weight would not require dermal data.

In addition, the experts of the EPAA project on acute toxicity are discussing ways of improving the current animal-based methods to support classification and labelling decisions in the agrochemical and chemical sectors. Among the opportunities considered, the replacement of lethality as an endpoint with appearance of clinical signs and the use of data from e.g. dose range finding studies, could contribute to reduce, refine and/or replace the current methods while maintaining human safety, as recognised during a workshop held in 2012 (EPAA, 2012).

6.4 Conclusions

Information on the acute systemic toxicity of chemicals is required for human health risk assessments under REACH and the Biocides Products Regulation, and can be used for classification and labelling under CLP.

Non-standard methods and ITS for acute systemic toxicity should predict LD_{50} values or official acute toxicity categories, depending on the regulatory application. Ideally, they should eventually be predictive of acute toxicity to humans.

Despite considerable research efforts over the past 20 years in the area of acute systemic toxicity, a complete mechanistic understanding of the key pathways is lacking, which is hampering the development of non-standard methods and AOP-based ITS. The ongoing EU FP7 SEURAT-1 research initiative, although focused on alternatives for repeat dose toxicity testing, is contributing to our understanding of toxicological modes of action, and could also provide innovative methodologies and tools for acute toxicity testing. The importance of biokinetic factors is well recognised, and yet there is still a need to further evaluate the usefulness of kinetic parameters in the context of ITS. The coupling of PBBK models with toxicodynamic models, as a means of predicting toxicologically relevant target tissue/organ doses *in vivo*, is currently being demonstrated in the EU FP7 COSMOS project. However, such models are available for relatively few chemicals, and thus do not support routine safety assessments.

Based on the current status of non-standard methods, it seems reasonable to use basal cytotoxicity methods, especially those that have been validated, for the identification of negatives (non-classified substances). Due to the limitations of these methods, results should always be used in combination with other information sources to build confidence

in the decision not to classify a substance for acute oral toxicity (EURL ECVAM, 2013). For the classification of acutely toxic substances, *in vitro* methods and QSARs may be suitable, but this would have to be evaluated on a case-by-case basis. To reduce animal testing for non-oral routes of exposure, extrapolation from *in vivo* oral toxicity to dermal toxicity is generally expected to be protective. However, the ability to extrapolate from the oral to the inhalational routes is less well established.

6.5 References

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485

Andrew D & Wright-Williams S (2011). Analysis of acute toxicity and irritation data submitted under the first REACH deadline. TSGE Poster presented at 2011 Autumn meeting of the British Toxicology Society.

Blaauboer BJ, Boekelheide K, Clewell HJ, Daneshian M, Dingemans MM, Goldberg AM, Heneweer M., Jaworska, J., Kramer NI, Leist M, Seibert H, Testai E, Vandebriel RJ, Yager JD & Zurlo J (2012). The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. *ALTEX* 29(4): 411-25.

Boekelheide K & Schuppe-Koistinen I (2012). SOT/EUROTOX debate: biomarkers from blood and urine will replace traditional histopathological evaluation to determine adverse responses. *Toxicol Sci.* 129(2): 249-55.

Bulgheroni A., Kinsner-Ovaskainen A., Hoffmann S, Hartung T & Prieto P (2009). Estimation of acute oral toxicity using the No Adverse Effect Level (NOAEL) from the 28-day repeated dose toxicity studies in rats. *Regulatory Toxicology and Pharmacology* 53: 16-19.

Chapman K, Creton S, Kupferschmidt H, Bond GR, Wilks MF & Robinson S (2010). The value of acute toxicity studies to support the clinical management of overdose and poisoning: a cross-discipline consensus. *Regul. Toxicol. Pharmacol.* 58(3): 354-9.

Coecke S, Pelkonen O, Leite SB, Bernauer U, Bessems JG, Bois FY, Gundert-Remy U, Loizou G, Testai E & Zaldivar JM (2013). Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicology in Vitro* 27(5): 1570-1577.

Creton S, Dewhurst IC, Earl LK, Gehen SC, Guest RL, Hotchkiss JA, Indans I, Woolhiser MR & Billington R. (2010). Acute toxicity testing of chemicals – Opportunities to avoid redundant testing and use alternative approaches. *Critical Reviews in Toxicology* 40(1): 50-83.

Ekwall B (1999). Overview of the Final MEIC Results: II. The In Vitro--In Vivo Evaluation, Including the Selection of a Practical Battery of Cell Tests for Prediction of Acute Lethal Blood Concentrations in Humans. *Toxicology in Vitro* 13(4-5): 665-673.

EURL ECVAM (2013). EURL ECVAM recommendation on the 3T3 Neutral Red Uptake (3T3 NRU) Cytotoxicity Assay for the Identification of Substances not Requiring Classification for Acute Oral Toxicity. Available from:

http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations

European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L396, 1-849.

European Commission (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353.

European Commission (2009a). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342, 59-209.

European Commission (2009b). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-47.

European Commission (2013). Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union L93, 1-84.

EPAA (2010). The European Partnership for Alternative Approaches to Animal Testing - Annual Report 2010. Available from: Available from: <https://circabc.europa.eu/sd/a/a219a8a8-3473-4fce-96ad-b8f50e2227a3/ar-2010.pdf> (accessed on 9 July 2014).

EPAA (2012). The European Partnership for Alternative Approaches to Animal Testing - Annual Report 2012. <https://circabc.europa.eu/sd/a/436f1b08-f36f-41b8-9058-912cbbd2cbb3/ar-2012.pdf> (accessed on 9 July 2014).

European Union (2008). Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L142

European Union (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.

European Union (2014). Commission Regulation (EU) No 260/2014 of 24 January 2014 amending, for the purpose of its adaptation to technical progress, Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L 81, 1-253.

Gennari A, van den Berghe C, Casati S, Castell J, et al. (2004). Strategies to replace in vivo acute systemic toxicity testing. The report and recommendations of ECVAM Workshop 50. ECVAM Workshop 50. *Alternatives to Laboratory Animals* 32(4): 437-459.

ICH (2009). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2). Recommended for Adoption at Step 4 of the ICH Process on June 11, 2009.

Indans I, Fry T, Parsons P & Evans P (1998). Classification and labelling of new industrial chemicals for acute toxicity, skin and eye irritation. *Human & Experimental Toxicology* 17: abstract 529.

Kinsner-Ovaskainen A, Prieto P, Stanzel S & Kopp-Schneider A (2013). Selection of test methods to be included in a testing strategy to predict acute oral toxicity: an approach based on statistical analysis of data collected in phase 1 of the ACuteTox project. *Toxicology in Vitro* 27(4): 1377-1394.

Lapenna S, Fuart-Gatnik M & Worth A (2010). Review of QSAR Models and Software Tools for predicting Acute and Chronic Systemic Toxicity. JRC Scientific and Technical Report. EUR 24639 EN. Available from: <http://publications.jrc.ec.europa.eu/repository/>

Liu LY, Yang T, Ji J, Wen Q, Morgan AA, Jin B, Chen G, Lyell DJ, Stevenson DK, Ling XB, Butte AJ., (2013). Integrating multiple 'omics' analyses identifies serological protein biomarkers for preeclampsia. *BMC Med.* 11(1): 236.

NIH (2006). Background Review Document (BRD): Validation of Neutral Red Uptake Test Methods NIH /In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity. National Institutes of Health Publication No. 07-4518, November 2006. Available from: http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm

NIH (2009). The report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing In Vitro Approaches and Humane Endpoints for Systemic Toxicity Evaluations. National Institutes of Health Available from: http://iccvam.niehs.nih.gov/docs/acutetox_docs/Workshop-Feb09/AcuteWkspRpt.pdf

Moore NP, Andrew DJ, Bjerke DL, Creton S, Dreher D, Holmes T, Prieto P, Seidle T & Rowan TG (2013). Can acute dermal systemic toxicity tests be replaced with oral tests? A comparison of route-specific systemic toxicity and hazard classifications under the Globally Harmonized System of Classification and Labelling of Chemicals (GHS). *Regulatory Toxicology and Pharmacology* 66(1): 30-37.

Norlén H, Berggren E, Whelan M & Worth A (2012). An investigation into the use of computational and in vitro methods for acute systemic toxicity prediction. JRC Scientific and Technical Report. EUR 25473 EN. Available from: <http://publications.jrc.ec.europa.eu/repository/>

OECD (1987). Guideline for Testing of Chemicals 402, Acute Dermal Toxicity. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070585-en.

OECD (2000). Guidance document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Series on Testing and Assessment N°19. Organisation for Economic Cooperation and Development Paris, France. Available from: <http://www.oecd.org>

OECD (2001a). Guideline for Testing of Chemicals, 420, Acute Oral Toxicity – Fixed Dose Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070943-en.

OECD (2001b). Guideline for Testing of Chemicals, 423, Acute Oral Toxicity – Acute Toxic Class Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071001-en.

OECD (2001c). Guideline for Testing of Chemicals, 425, Acute Oral Toxicity – Up-and-Down Procedure. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071049-en.

OECD (2009a). Guideline for Testing of Chemicals, 403, Acute Inhalation Toxicity. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070608-en

OECD (2009b). Guideline for Testing of Chemicals, 436, Acute Inhalation Toxicity – Acute Toxic Class Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076037-en

OECD (2010). Guidance Document No 129 on using Cytotoxicity Tests to Estimate Starting Doses for Acute Oral Systemic Toxicity Tests. Organisation for Economic Cooperation and Development Paris, France. Available from: <http://www.oecd.org>

Prieto P, Kinsner-Ovaskainen A, Stanzel, et al. (2013a). The value of selected in vitro and in silico methods to predict acute oral toxicity in a regulatory context: Results from the European Project ACuteTox. *Toxicology in Vitro* 27(4): 357-376.

Prieto P, Cole T, Curren R, Gibson RM, Liebsch M, Raabe H, Tuomainen AM, Whelan M, Kinsner-Ovaskainen A (2013b). Assessment of the predictive capacity of the 3T3 neutral red uptake cytotoxicity test method to identify substances not classified for acute oral toxicity (LD50 > 2000 mg/kg): results of an ECVAM validation study. *Regulatory Toxicology and Pharmacology* 65(3), 344-65.

Robinson S, Delongea JL, Donald E, Dreher D, Festag M, Kervyn S, Lampo A, Nahas K, Nogues V, Ockert D, Quinn K, Old S, Pickersgill N, Somers K, Stark C, Stei P, Waterson L & Chapman K (2008). A European pharmaceutical company initiative challenging the regulatory requirement for acute toxicity studies in pharmaceutical drug development. *Regulatory Toxicology and Pharmacology* 50: 345-352.

Seidle T, Robinson S, Holmes T, Creton S, Prieto P, Scheel J & Chlebus M (2010). Cross-sector review of drivers and available 3Rs approaches for acute systemic toxicity testing. *Toxicological Sciences* 116(2): 382-396.

Seidle T, Prieto P & Bulgheroni A (2011). Examining the Regulatory Value of Multi-route Mammalian Acute Systemic Toxicity Studies. *Alternatives to Animal Experimentation* 28(2): 95-102.

Schrage A, Hempel K, Schulz M, Kollé SN, van Ravenzwaay B & Landsiedel R (2011). Refinement and Reduction of Acute Oral Toxicity Testing: A Critical Review of the Use of Cytotoxicity Data. *Alternatives to Laboratory Animals* 39: 273-295.

Thomas HD & Dewhurst IC (2007). What Does a Dermal Acute Toxicity Study Add to the Information on a Plant Protection Pesticide? *Abstracts Toxicology* 231: 114-115.

Zurich MG, Stanzel S, Kopp-Schneider A, Prieto P & Honegger P (2013). Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity. *Toxicology in Vitro* 27(4): 1416-1424.

Chapter 6. Table of References with Notes

Traditional animal tests – OECD TGs and GD	
<p>Organisation for Economic Cooperation and Development (OECD). 1987. Guideline for Testing of Chemicals, 402, Acute Dermal toxicity. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070585-en.</p>	<p>Dermal route. This method is based on the classical LD50 test.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2000. Guidance document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Series on Testing and Assessment N°19. Paris, France. Available from: http://www.oecd.org.</p>	<p>The purpose of this Guidance Document is to apply the principles of the Three Rs to the use of animals in regulatory toxicity tests. It specifically addresses Refinement. The objective of the document is to provide useful guidance and criteria for determining when an animal is in a moribund condition, or expected to become moribund, or experiencing significant pain and distress, and should therefore be euthanised. The ultimate purpose of the application of humane endpoints to toxicology studies is to be able to accurately predict severe pain, severe distress, suffering, or impending death, before the animal experiences these effects.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2001a. Guideline for Testing of Chemicals, 420, Acute Oral Toxicity – Fixed Dose Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070943-en.</p> <p>EU Test Method B.1 bis. Acute Oral Toxicity — Fixed Dose Procedure</p>	<p>Oral route. This method is based on a stepwise approach. It does not use lethality as endpoint but evident signs of toxicity. It provides an estimated LD50 range.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2001b. Guideline for Testing of Chemicals, 423, Acute Oral Toxicity – Acute Toxic Class Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071001-en.</p> <p>EU Test Method B.1 tris. Acute Oral Toxicity — Acute Toxic Class Method</p>	<p>Oral route. This method is based on a stepwise approach. It uses lethality as endpoint. It provides an estimated LD50 range.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2001c. Guideline for Testing of Chemicals, 425, Acute Oral Toxicity – Up-and-Down Procedure. OECD Guidelines for the Testing of Chemicals, Section 4, OECD</p>	<p>Oral route. This method is based on a stepwise approach. It uses lethality as endpoint. It provides LD50 point estimate with confidence intervals.</p>

Publishing. doi: 10.1787/9789264071049-en.	
<p>Organisation for Economic Cooperation and Development (OECD). 2009a. Guideline for Testing of Chemicals, 403, Acute Inhalation Toxicity. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070608-en</p> <p>EU Test Method B.2 Acute Inhalation Toxicity</p>	<p>Inhalation route. This revised Test Guideline describes the traditional LC50 protocol and a Concentration x Time (C x t) protocol. It can be used to estimate a median lethal concentration (LC50), non-lethal threshold concentration (LC01) and slope, and to identify possible sex susceptibility</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2009b. Guideline for Testing of Chemicals, 436, Acute Inhalation Toxicity – Acute Toxic Class Method. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264076037-en.</p> <p>EU Test Method B.52. Acute Inhalation Toxicity — Acute Toxic Class Method</p>	<p>Inhalation route. The test method is based on a stepwise procedure, each step using 3 animals of each sex (the preferred species is rat). Animals are exposed in inhalation chambers to a pre-defined concentration for 4 hours.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2010. Guidance Document No 129 on using Cytotoxicity Tests to Estimate Starting Doses for Acute Oral Systemic Toxicity Tests. Paris, France. Available from: http://www.oecd.org</p>	<p>This guidance document describes methods to determine the in vitro basal cytotoxicity of test substances and to use these in vitro data to estimate the starting dose for the acute oral in vivo test.</p>
Regulatory requirements across sectors	
<p>European Commission, 2006. Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L396, 1-849.</p>	<p><i>Chemicals:</i> Information requirements depend on the production volume of the chemical concerned:</p> <ul style="list-style-type: none"> • Annex VII (1-10 tpa): acute toxicity by the oral route. <ul style="list-style-type: none"> ○ The study/ies do(es) not generally need to be conducted if the substance is classified as corrosive to the skin or if a study on acute toxicity by the inhalation route is available. • Annex VIII-X (10->1000 tpa): In addition to the oral route, for substances other than gases, the information shall be provided for at least one other route. The choice for the second route will depend on the nature of the substance and the likely route of human exposure. If there is only one route of exposure, information for only that route need to be provided. <ul style="list-style-type: none"> ○ Testing by the inhalation route is appropriate if exposure of humans via inhalation is likely taking into account the vapour pressure of the substance and/or the possibility of exposure to aerosols, particles or droplets of an inhalable size.

	<ul style="list-style-type: none"> ○ Testing by the dermal route is appropriate if: (1) inhalation of the substance is unlikely; and (2) skin contact in production and/or use is likely; and (3) the physicochemical and toxicological properties suggest potential for a significant rate of absorption through the skin.
<p>European Union, 2008. Council Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) AND</p> <p>European Union, 2014. Commission Regulation (EU) No 260/2014 of 24 January 2014 amending, for the purpose of its adaptation to technical progress, Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L 81, 1-253</p>	<p>The Regulation of May 2008 and its subsequent amendments, define tests testing of chemicals for the REACH Regulation. They are based on the OECD Guidelines for the Testing of Chemicals.</p> <p>The January 2014 adaptation contains, among others, nine methods for the determination of toxicity and other health effects including four inhalation toxicity test methods, which include an update of three methods and one new method to reduce the number of animals used and to improve assessment of effects, an update of the repeat dose 28-day oral toxicity test method to include parameters for assessment of endocrine activity, an update of the toxicokinetics test method relevant for the design and understanding of toxicological studies and an update of chronic, carcinogenicity and combined chronic and carcinogenicity test methods.</p>
<p>European Commission, 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353.</p>	<p><i>EU CLP Regulation</i></p> <p>Substances can be allocated to one of four toxicity categories based on acute toxicity by the oral, dermal or inhalation route according to the numeric criteria.</p>
<p>Robinson, S., Delongas, J.L., Donald, E., Dreher, D., Festag, M., Kervyn, S., Lampo, A., Nahas, K., Nogues, V., Ockert, D., Quinn, K., Old, S., Pickersgill, N., Somers, K., Stark, C., Stei, P., Waterson, L., Chapman, K., 2008. A European pharmaceutical company initiative challenging the regulatory requirement for acute toxicity studies in pharmaceutical drug development. Regul. Toxicol. Pharmacol. 50, 345-352.</p>	<p>This paper describes the results of an evidence-based review of acute toxicity studies undertaken by 13 European pharmaceutical companies and 3 research organisations, aimed to assess the value of the data generated. The two main objectives were to review how acute toxicity data were gathered and used across the pharmaceutical industry and to develop a strategy for challenging the guidelines on the requirement for acute toxicity where lethality is an endpoint. The ultimate objective was to assess the usefulness of acute toxicity in the drug development process and to provide recommendations for alternative strategies that use information derived from other studies. The group concluded that acute toxicity studies are not needed prior to first clinical trial in humans. As a direct result of this work, the requirement for acute toxicity data prior to first in man clinical trials has been removed from ICH M3 guidelines.</p>
<p>European Commission, 2009a. Regulation (EC) No 1223/2009</p>	<p><i>Cosmetics</i>: This Regulation establishes rules to be complied with by any cosmetic product made</p>

<p>of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342, 59-209.</p>	<p>available on the market, in order to ensure the functioning of the internal market and a high level of protection of human health. Article 18 (Animal testing)</p>
<p>European Commission, 2009b. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-47. AND Commission Regulation, 2013. Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market</p>	<p><i>Plant Protection Products:</i></p> <ul style="list-style-type: none"> • The acute oral toxicity of the active substance shall always be reported. • The acute dermal toxicity of the active substance shall be reported unless waiving is scientifically justified (for example where oral LD₅₀ is greater than 2000 mg/kg). Both local and systemic effects shall be investigated. • The acute inhalation toxicity of the active substance shall be reported where any of the following apply: <ul style="list-style-type: none"> ○ the active substance has a vapour pressure > 1 × 10⁻² Pa at 20 °C; ○ the active substance is a powder containing a significant proportion of particles of a diameter < 50 µm (> 1 % on weight basis); ○ the active substance is included in products that are powders or are applied by spraying
<p>ICH, 2009. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2). Recommended for Adoption at Step 4 of the ICH Process on June 11, 2009.</p>	<p><i>Pharmaceuticals:</i> Acute toxicity data prior to first in man clinical trials are not any longer required. When the acute toxicity information is available from any study, separate single-dose studies are not recommended. Studies providing acute toxicity information can be limited to the clinical route only and such data can be obtained from non-GLP studies if clinical administration is supported by appropriate GLP repeated-dose toxicity studies. Lethality should not be an intended endpoint in studies assessing acute toxicity.</p>
<p>Chapman, K., Creton, S., Kupferschmidt, H., Bond, G.R., Wilks, M.F., Robinson, S., 2010. The value of acute toxicity studies to support the clinical management of overdose and poisoning: a cross-discipline consensus. Regul. Toxicol. Pharmacol. 58(3):354-9.</p>	<p>This paper describes the outcome of a workshop aimed to explore whether data from acute toxicity studies are used by clinicians and Poison Centres to assess and treat human overdoses and poisoning. The key conclusions were the following:</p> <ul style="list-style-type: none"> • Acute toxicity studies do not provide information to clinicians on how to treat poisoning and overdose patients for pharmaceutical and non-pharmaceutical chemicals. • Human data are of most value to Poison Centres in informing treatment of overdose and poisoning. • More useful animal toxicity data could be obtained from nonlethal studies that are already carried out as part of the product development process. • In sectors where acute toxicity may be the only systemic toxicity study performed, the value of such studies could be maximised by improving the design. • There is little differentiation in hazard statements and treatment advice between classification and labelling categories and therefore this is of no value in the specific management of

	<p>poisoning.</p> <p>The consensus formed at the workshop was that acute toxicity studies are not used for managing overdose of pharmaceuticals and are of little value to treat human poisoning from chemicals. This was the last remaining driver for acute toxicity studies for pharmaceuticals and it has been removed.</p>
<p>European Union, 2012. Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.</p>	<p><i>Biocides</i>: Annex II, Data Requirement 8.7 (acute toxicity): in addition to the oral route testing via other routes is only necessary when specified criteria are met.</p> <p>for chemical Active Substances: ‘<i>Before a new dermal acute toxicity study is carried out, an in vitro dermal penetration study (OECD 428) should be conducted to assess the likely magnitude and rate of dermal bioavailability</i>’.</p> <p>Testing by the dermal route is necessary only if: ‘<i>inhalation of the substance is unlikely; or skin contact in production and/or use is likely; and either the physicochemical and toxicological properties suggest potential for a significant rate of absorption through the skin; or the results of an in vitro dermal penetration study (OECD 428) demonstrate high dermal absorption and bioavailability</i>’.</p>
<p>Mechanistic understanding</p>	
<p>Ekwall B., 1999. Overview of the Final MEIC Results: II. The In Vitro--In Vivo Evaluation, Including the Selection of a Practical Battery of Cell Tests for Prediction of Acute Lethal Blood Concentrations in Humans. <i>Toxicology In Vitro</i> 13 (4-5):665-73.</p>	<p>This paper provides an overview of the results of the Multicenter Evaluation of <i>In Vitro</i> Cytotoxicity 7-year programme that was set up by the Scandinavian Society of Cell Toxicology with the main goal of evaluating the relevance of <i>in vitro</i> toxicity assays for human acute toxicity. Fifty chemicals were tested in 29 laboratories in 61 cytotoxicity assays. The basal cytotoxicity hypothesis claims that a) a majority of chemicals cause lethal toxicity in humans by interfering with functions common to all human cells, b) basal cytotoxicity can be tested in undifferentiated cell lines (preferably of human origin), c) only a fraction of the IC₅₀ for cell lines will lead to human lethality through functional interference with vital neurons in the brain, d) the basal cytotoxicity is a baseline minimal toxicity for all chemicals – thus, cell tests will not give false positive results. The correlative/mechanistic <i>in vitro/in vivo</i> evaluation of the results supported the four hypotheses.</p>
<p>Gennari A, van den Berghe C, Casati S, Castell J, et al., 2004. Strategies to replace in vivo acute systemic toxicity testing. The report and recommendations of ECVAM Workshop 50. <i>ECVAM Workshop 50. Altern Lab Anim.</i> 2004 32(4):437-59.</p>	<p>This report reviews the state-of-the-art of <i>in vitro</i> methods for estimating acute systemic toxicity and to develop research, development and validation strategies necessary for the replacement of <i>in vivo</i> testing with a focus on the oral route of exposure. The workshop participants developed a list of potential susceptible targets that are common to many cell types and that could serve as <i>in vitro</i> endpoints for acute toxicity. Kidney, liver, central nervous system, cardiovascular system, lung/respiratory system, blood and the gastrointestinal tract were identified as the main organ systems known to be susceptible to acute toxicity resulting in lethality. The susceptible function for each organ was also identified. An attempt was also made to identify the common</p>

	<p>physiological processes and the assay systems for quantifying specific endpoints related to each of them. The main outcome of the workshop was to suggest the development of an integrated testing strategy based on the use of physicochemical properties, <i>in vitro-in vivo</i> data, computational methods, basal cytotoxicity assays and complementary assays for metabolisms, transport, kinetics and target organ toxicity.</p>
<p>National Institutes of Health (NIH), 2009. The report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing In Vitro Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Available from: http://iccvam.niehs.nih.gov/docs/acutetox_docs/Workshop-Feb09/AcuteWkspRpt.pdf</p>	<p>The development of alternative test methods and subsequent replacement of animals in acute oral toxicity testing calls for a better understanding of critical toxicity pathways. The scientific workshop on acute chemical safety testing contributed to this proof-of-concept by developing approaches to identify the key toxicity pathways for acute systemic toxicity. The participants discussed (a) the current understanding of key pathways for <i>in vivo</i> acute systemic toxicity and identified knowledge gaps, (b) the current acute systemic toxicity injury and toxicity assessment, (c) earlier humane endpoints for acute systemic toxicity, (d) the application of <i>in vitro</i> mode of action and mechanistic information to the development and validation of <i>in vitro</i> methods for assessing acute systemic toxicity, and (e) industry involvement in test method development, validation and use. The workshop suggested that the mechanistic information could be used to develop predictive <i>in vitro</i> alternative test methods and might also help to identify predictive biomarkers of systemic toxicity for use as earlier, more humane endpoints during <i>in vivo</i> tests, thereby reducing pain and distress.</p>
<p>Boekelheide K & Schuppe-Koistinen I. 2012. SOT/EUROTOX debate: biomarkers from blood and urine will replace traditional histopathological evaluation to determine adverse responses. Toxicol Sci. 129(2): 249-55.</p>	<p>The 2011 SOT/EUROTOX debate addressed the proposition that “Biomarkers From Blood and Urine will Replace Traditional Histopathological Evaluation to Determine Adverse Responses,” identifying and comparing the strengths and limitations of histopathology and serum and urine biomarkers. This article discusses the uses and the limitations of having a gold standard, how adverse responses are determined, the evolutionary (as opposed to revolutionary) process by which one technology is typically replaced by another, and the overall goal of developing biomarkers which can translate from preclinical safety assessment to clinical utility. The ultimate purpose is to help researchers and regulators understand the challenges they face in the development and integration of new and existing biomarkers to determine adverse responses.</p> <ul style="list-style-type: none"> • Biomarkers have been used in clinical diagnosis and drug development for decades to measure physiological parameters such as blood pressure, heart rate, body temperature, and urine color, and clinical chemistry parameters such as enzymes in blood and urine, serum creatinine levels, and urinary glucose. • Advances in molecular biology and analytical technologies have enabled the discovery of novel biomarkers which include individual genes, transcripts, proteins, and endogenous metabolites (including lipids), and their associated patterns of expression. Other examples include the use of imaging patterns, electrical signals, and cell levels in body fluids to identify

	<p>potential concern.</p> <ul style="list-style-type: none"> • Today, it is technically feasible to identify tissue-specific markers through the use of these – omics technologies by integrating and correlating toxicology data. These changes in gene, protein, and metabolite expressions are consistent, sensitive, and early, and provide the molecular basis of drug-induced injuries. The genomic-derived candidate markers can then be localized to organs using in situ hybridization and immunohistochemistry for their proteins. The perturbations at the transcript level are mirrored on the protein level and subsequently secreted into body fluids where they can be measured. • The use of safety biomarkers is increasing the confidence of decision making in drug research and development; early-stage projects can be pulled from development more quickly, and late-stage projects become less risky when translational safety biomarker data provide the clarity, the predictability, and the possibility to monitor drug-induced toxicity. New technologies and molecular approaches, including molecular profiling and molecular pathology as a complement to the classical toolbox, promise a future of further discoveries and improvements in the safety biomarker field.
<p>Liu LY, Yang T, Ji J, Wen Q, Morgan AA, Jin B, Chen G, Lyell DJ, Stevenson DK, Ling XB, Butte AJ., 2013. Integrating multiple 'omics' analyses identifies serological protein biomarkers for preeclampsia. BMC Med. 11(1): 236.</p>	<p>The authors have employed a comprehensive unbiased multi-'omics' approach, integrating results from microarray multiplex analysis and proteomic identification by two-dimensional gel analysis with the ultimate aim of discovering diagnostic biomarkers for preeclampsia (PE), a multifactorial disease for which it is likely unrealistic that a single biomarker could be used for diagnoses. They have used a novel concept: combining a transcriptomic approach in placenta tissue (closer to the focus of the pathophysiology) with a proteomic approach in serum (which is more appropriate for clinical use). In addition to the previously identified biomarkers, the authors found 3 up-regulated and 6 down-regulated biomarkers in PE sera. Two optimal biomarker panels were developed for early and late onset PE assessment, respectively. The functional significance of these PE biomarkers and their associated pathways were also analysed, which may provide new insights into the pathogenesis of PE.</p>
<p>Biokinetic considerations</p>	
<p>Adler, S., Basketter, D., Creton, S., et al., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch. Toxicol. 85(5):367-485.</p>	<p>This report provides an analysis made by a panel of experts of the status and prospects of alternative methods in five toxicological areas of concern in view of the full marketing ban foreseen for 2013 for cosmetics products and ingredients tested on animals in Europe. The report includes five chapters dealing with toxicokinetics, skin sensitisation, repeated dose toxicity, carcinogenicity and reproductive toxicity.</p> <p>Information on toxicokinetics is indispensable for the development and design of more efficient strategies, for <i>in vitro-in vivo</i> extrapolation, and for the identification of clearance and the role of metabolites.</p>

<p>Blaauboer, B.J., Boekelheide, K., Clewell, H.J., Daneshian, M., Dingemans, M.M., Goldberg, A.M., Heneweer, M., Jaworska, J., Kramer, N.I., Leist, M., Seibert, H., Testai, E., Vandebriel, R.J., Yager, J.D., Zurlo, J., 2012. The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. ALTEX 29(4):411-25.</p>	<p>This report presents the results of a workshop aimed at better defining the use of in vitro-derived biomarkers of toxicity and determining the place these data can have in human risk assessment. As a result, a conceptual framework is presented for the incorporation of in vitro-derived toxicity data into the risk assessment process. The implementation of the proposed framework requires the collection and collation of data from existing literature and new <i>in vitro</i> test systems, as well as the categorization of biomarkers of toxicity and their relation to pathways-of-toxicity. Data selection and integration need to be also driven by their usefulness in a quantitative in vitro-in vivo extrapolation.</p>
<p>Coecke, S., Pelkonen, O., Leite, S.B., Bernauer, U., Bessems, J.G., Bois, F.Y., Gundert-Remy, U., Loizou, G., Testai, E., Zaldívar, J.M., 2013. Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. Toxicol. In Vitro 27(5):1570-7.</p>	<p>This paper discusses the importance of toxicokinetics as a key element to integrate the results from <i>in silico</i>, <i>in vitro</i> and already available <i>in vivo</i> studies, in particular when non-animal testing approaches need to be implemented. Toxicokinetics is needed to estimate the range of target organ doses that can be expected from realistic human external exposure scenarios. This information is crucial for determining the dose/concentration range that should be used for <i>in vitro</i> testing. Toxicokinetics is necessary to convert the <i>in vitro</i> results, generated at tissue/cell or sub-cellular level, into dose response or potency information relating to the entire target organism, i.e. the human body (<i>in vitro</i>–<i>in vivo</i> extrapolation, IVIVE).</p>
<p><i>In vitro</i> NRU cytotoxicity assays</p>	
<p>National Institutes of Health (NIH), 2006. Background Review Document (BRD): Validation of Neutral Red Uptake Test Methods NIH /In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity. Publication No. 07-4518, November 2006. Available from: http://icevam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm</p>	<p>This report provides an overview of the <i>in vitro</i> cytotoxicity data generated during the NICEATM/ECVAM In Vitro Cytotoxicity Validation Study to predict rodent <i>in vivo</i> LD₅₀ values and starting doses for acute oral systemic toxicity test methods. The <i>in vitro</i> tests evaluated used rodent (mouse fibroblast [3T3]) and human (normal human epidermal keratinocyte [NHK]) cells. The validation study had the following objectives: a) to further standardise and optimise the <i>in vitro</i> basal cytotoxicity protocols to maximize test reliability (intra- and inter-laboratory reproducibility); b) to assess the accuracy of the standardised <i>in vitro</i> cytotoxicity test methods for estimating rodent oral LD₅₀ values across the five UN GHS categories of acute oral systemic toxicity, as well as unclassified toxicities; c) to estimate the reduction and refinement in animal use achievable from using the <i>in vitro</i> basal cytotoxicity test methods to identify starting doses for <i>in vivo</i> acute oral systemic toxicity tests; d) to develop databases containing high quality data from <i>in vivo</i> acute oral lethality and <i>in vitro</i> basal cytotoxicity tests that can be used to support the investigation of other <i>in vitro</i> test methods necessary to improve the prediction of <i>in vivo</i> acute oral lethality.</p> <p>An independent scientific review panel evaluated the validation status of the 3T3 and NHK basal cytotoxicity test methods for use as adjuncts to <i>in vivo</i> acute oral systemic toxicity tests for the purpose of determining starting doses and concluded that they should be considered for use in a</p>

	<p>weight-of-evidence approach to determine starting doses for acute oral systemic toxicity test methods (i.e., the Up-and-Down Procedure and the Acute Toxic Class method) to reduce the use of animals required for these methods. The panel also concluded that these <i>in vitro</i> cytotoxicity test methods could not be used to determine the hazard classification of chemicals.</p>
<p>Schrage, A., Hempel, K., Schulz, M., Kolle, S.N., van Ravenzwaay, B., Landsiedel, R., 2011. Refinement and Reduction of Acute Oral Toxicity Testing: A Critical Review of the Use of Cytotoxicity Data. <i>Alt. Lab. Anim.</i> 39:273-295.</p>	<p>This paper shows an analysis aimed to estimate the value of the 3T3 NRU <i>in vitro</i> method for predicting the <i>in vivo</i> acute oral classification and for predicting the starting dose for the subsequent <i>in vivo</i> test. The authors also calculated the animal numbers required in each instance, in order to estimate possible reductions in animal use. Additionally, they tried to identify specific applicability domains, hoping to improve the predictivity of the cytotoxicity test. Two hundred and two substances (187 of which were BASF in-house substances) including a broad variety of chemicals, agrochemicals and formulations were tested. The overall capacity of the 3T3 NRU uptake to predict the EU-GHS acute oral toxicity categories was low (35%), while a good concordance of 74%, was obtained only for the weakly toxic substances (EU-GHS Cat. 4). The estimation of the starting dose by the cytotoxicity test was useful for 59% of the substances. However, the use of a standard starting dose of 300mg/kg b.w. by default would have been almost as useful (50%). In contrast, the prediction by an experienced toxicologist was correct for 95% of the substances. However, this was only performed for 40% of the substances, mainly those of no to low toxicity. The outcome supports the argument that 3T3 NRU cytotoxicity data cannot be used as stand alone to assess the acute oral toxicity of test substances.</p>
<p>EURL ECVAM recommendation on the 3T3 Neutral Red Uptake (3T3 NRU) Cytotoxicity Assay for the Identification of Substances not Requiring Classification for Acute Oral Toxicity. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations</p>	<p>The 3T3 NRU test method could prove a valuable component of a WoE or ITS approach for supporting hazard identification and safety assessment in agreement with the EU CLP Regulation and international regulatory schemes implementing the upper threshold of UN GHS Category 4 as the cut-off for non-classification of substances. In particular, data from the 3T3 NRU assay may constitute an information source within a WoE approach under the provisions of the REACH regulation (Annex XI, 1.2) potentially supporting conclusions on absence of acute oral toxicity of industrial chemicals.</p> <p>(1) Considering the results of the EURL ECVAM study and the data available from the previous NICEATM/EURL ECVAM validation study (NIH, 2006), the 3T3 NRU test method shows a high sensitivity (ca 95%) and, consequently, a low false negative rate (ca 5%) when employed in conjunction with a prediction model to distinguish potentially toxic versus non-toxic (i.e. classified versus non-classified) substances. Therefore, substances found to be negative in this test would most likely not require classification for acute oral toxicity based on a cut-off value of >2000 mg / kg b.w.. Concluding on likely non-classification however requires careful consideration of the limitations of the assay (see points 2 and 3).</p>

- (2) The 3T3 NRU test method is sensitive to hazardous chemicals acting through general mechanisms of toxicity common to most cell types, often referred to as 'basal cytotoxicity'. However, chemicals not exhibiting significant cytotoxicity but which act through mechanisms specific only to certain cell types and tissues (e.g. of the heart or central nervous system) may not be indicated as potentially acutely toxic by this method. Moreover, chemicals requiring metabolic activation to induce toxicity may also go undetected since the cell model lacks significant metabolic capacity. Care must be taken therefore in interpreting negative results derived from this assay, despite the low false negative rate demonstrated in the EURL ECVAM validation study.
- (3) Considering its limitations, results derived from the 3T3 NRU test method should be always used in combination with other information sources to build confidence in the decision not to classify a substance for acute oral toxicity. Possible complementary information sources include chemical analogues, physico-chemical properties, structural alerts, structure–activity relationships, and toxicokinetic data. The 3T3 NRU method would fit well therefore within a Weight of Evidence (WoE) approach or as a component of an Integrated Testing Strategy (ITS).
- (4) The 3T3 NRU test method appears to be particularly relevant for the assessment of industrial chemicals since they are not designed to act on specific biological targets and in general tend not to be acutely toxic. Moreover, industrial chemicals which do exhibit toxicity are likely to act through multiple non-specific mechanisms that affect most cell types (i.e. basal cytotoxicity). Following the provisions of the REACH Regulation (1907/2006/EU) and in particular its Annex XI, data from the 3T3 NRU test method could be used within a WoE approach to adapt the standard information requirements.
- (5) The 3T3 NRU test method has a high false positive rate therefore positive results cannot be readily used in a meaningful way. A likely reason is that the test method does not capture important biokinetic processes such as absorption, distribution, metabolism and excretion. Thus, certain chemicals, despite having cytotoxic potential, may not actually be acutely toxic via the oral route.
- (6) As this test method informs about basal cytotoxicity which is a key event in many prevalent toxicological modes-of-action associated with both acute and chronic health effects, and since the EURL ECVAM validation study has shown the method to be amenable to automation and High Throughput Screening (HTS), it may constitute a valuable and economical information source for hazard profiling of substances.
- (7) Respecting the provisions of Directive 2010/63/EU on the protection of animals used for scientific purposes, before embarking on animal experiments to identify acute oral toxicity the 3T3 NRU test method should be considered as an initial screening tool together with complimentary information in order to possibly reduce or avoid animal testing.

<p>Prieto, P., Curren, R., Gibson, R.M., Raabe, H., Tuomainen, A.M., Whelan, M., Kinsner-Ovaskainen, A., 2013b. Assessment of the predictive capacity of the 3T3 neutral red uptake cytotoxicity test method to identify substances not classified for acute oral toxicity (LD₅₀ > 2000 mg/kg): results of an ECVAM validation study, <i>Regulatory Toxicology and Pharmacology</i> 65(3): 344-365.</p>	<p>This paper describes the results of a validation study that was organised to assess if the 3T3 Neutral Red Uptake cytotoxicity assay could identify substances not requiring classification as acute oral toxicants under the EU regulations. Fifty six coded industrial chemicals were tested. The assay exhibited high sensitivity (92-96%) but relatively low specificity (40-44%). The false negative rate was very low (ca 5%) and only three chemicals were under predicted. Assuming that most industrial chemicals are not likely to be acutely toxic, this test method could prove a valuable component of an integrated testing strategy, a read-across argument, or weight-of-evidence approach to identify non toxic chemicals (LD₅₀ > 2000 mg/kg). However, it is likely to under predict chemicals not exhibiting significant cytotoxicity but which act through mechanisms specific to certain cell types and tissues (e.g. of the heart or central nervous system) not captured by the 3T3 test system, or which first require biotransformation <i>in vivo</i>.</p>
<p>QSARs</p>	
<p>Lapenna S, Fuart-Gatnik M & Worth A (2010). Review of QSAR Models and Software Tools for predicting Acute and Chronic Systemic Toxicity. JRC Scientific and Technical Report. EUR 24639 EN. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This JRC report provides a review of different computational estimation methods for predicting acute and chronic systemic toxicity. It provides an overview of Quantitative Structure-Activity Relationship (QSAR) models published in the literature, commonly used software tools, and available databases suitable for QSAR analysis. It also briefly explains the Threshold of Toxicological Concern (TTC) concept and how this is used in prioritising chemicals for further assessment and preliminary risk characterisation.</p>
<p>Norlén H, Berggren E, Whelan M & Worth A (2013). An investigation into the use of computational and in vitro methods for acute systemic toxicity prediction. JRC Scientific and Technical Report. EUR 25473 EN. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This report describes a JRC study on the abilities of five alternative approaches to predict acute oral toxicity. The authors investigated four QSAR models (ToxSuite, TOPKAT, TEST and ADMET Predictor) and one in vitro method (3T3 NRU). Based on a test set of in vitro and in vivo data for 180 compounds, they characterized the predictive performance of each method when used alone (both for LD₅₀ prediction and acute toxicity classification into three categories), as well as multiple test combinations (batteries) and stepwise testing strategies (for acute toxicity classification into three categories). When used individually, the alternative methods showed an ability to predict LD₅₀ with correlation coefficients in the range from 49% to 84%, and to classify into three toxicity groups with accuracies in the range from 41% to 72%. When the alternative methods were combined into batteries or testing strategies, the overall accuracy of prediction could reach 76%. The authors also illustrate how different combinations of methods can be used to optimize sensitivity or specificity.</p>
<p>Testing strategies – Results from the EU FP6 Project ACuteTox aimed at the development of non-animal testing strategies for predicting human acute oral toxicity</p>	
<p>Kinsner-Ovaskainen, A., Prieto, P., Stanzel, S., Kopp-</p>	<p>This paper describes the methods used for statistical evaluation of concentration-response data</p>

<p>Schneider, A. 2013. Selection of test methods to be included in a testing strategy to predict acute oral toxicity: an approach based on statistical analysis of data collected in phase 1 of the ACuteTox project. <i>Toxicology In Vitro</i> 27(4): 1377-1394</p>	<p>collected for each of the endpoint assays used in the first phase of the ACuteTox project (assessing specific organ- and system-toxicity, such as haemato-, neuro-, nephro- and hepatotoxicity, as well as intestinal absorption, distribution and metabolism), and for the development of a testing strategy applicable for acute oral toxicity classification of chemicals based on the achieved results of the concentration-response analysis. The final list of <i>in vitro</i> test methods considered to be promising candidates for building blocks of the testing strategy is presented.</p>
<p>Prieto, P., Kinsner-Ovaskainen, A., Stanzel, et al., 2013a. The value of selected <i>in vitro</i> and <i>in silico</i> methods to predict acute oral toxicity in a regulatory context: Results from the European Project ACuteTox. <i>Toxicol In Vitro</i> 27(4): 357-376.</p>	<p>This paper describes the classification approaches studied in the ACuteTox project (single step procedures and two step tiered testing strategies) to evaluate the capacity of several competing non-animal testing strategies to correctly classify compounds into the four EU CLP acute oral toxicity categories and the non-classified. In summary, four <i>in vitro</i> testing strategies were proposed as best performing. In addition, a heuristic testing strategy is suggested that combines the prediction results gained from the 3T3 NRU cytotoxicity assay, with information on neurotoxicity alerts identified by the primary rat brain aggregates test method. Octanol–water partition coefficients and <i>in silico</i> prediction of intestinal absorption and blood–brain barrier passage were also considered. This approach allowed to reduce the number of chemicals wrongly predicted as non-classified (oral LD₅₀ > 2000 mg/kg b.w.).</p>
<p>Zurich MG, Stanzel S, Kopp-Schneider A, Prieto P, Honegger P., 2013. Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity. <i>Toxicol. In Vitro</i>, 27(4): 1416–1424</p>	<p>Primary aggregating brain cell cultures (AGGR) were examined for their capability to detect organ-specific toxicity apparently missed by using the 3T3 NRU <i>in vitro</i> cytotoxicity assay. The lowest observed effect concentration determined for each chemical (86 chemicals tested) was compared with the IC₂₀ reported for the 3T3 NRU cytotoxicity assay. The results showed that the frequency of alerts increased with the level of toxicity observed in AGGR. The overall findings suggest that AGGR are suitable for the detection of organ-specific toxicity and that they could, in conjunction with the 3T3 NRU cytotoxicity assay, improve the predictive capacity of <i>in vitro</i> toxicity testing.</p>
<p>Estimation of not-classified chemicals from repeat dose studies</p>	
<p>Bulgheroni, A., Kinsner-Ovaskainen, A., Hoffmann, S., Hartung, T., Prieto, P., 2009. Estimation of acute oral toxicity using the No Adverse Effect Level (NOAEL) from the 28-day repeated dose toxicity studies in rats. <i>Regulatory Toxicology and Pharmacology</i> 53: 16-19.</p>	<p>This paper proposes an approach to identify non toxic compounds (oral LD₅₀ > 2000 mg/kg) using information from 28 days repeated dose toxicity studies. The data used in this study were retrieved from the New Chemical Database (NCD). This database was accessed on 27th March 2008. Of 7200 notifications since 1981, 4773 substances were found of which 4219 included oral toxicity data. The prevalence of the different oral acute toxicity classes was analysed within the new chemicals notified in the NCD and a high prevalence of non-toxic substances (87%) was found. It was possible to set a NOAEL threshold of ≥ 200 mg/kg that allowed the correct identification of 63% of non-toxic compounds, while <1% of harmful compounds were misclassified as non-toxic. This approach could be useful under REACH and other chemical assessment programs where the 28-day study is legally required.</p>

Redundant testing – waiving possibilities	
Indans, I., Fry, T., Parsons, P., Evans, P., 1998. Classification and labelling of new industrial chemicals for acute toxicity, skin and eye irritation. Hum Exp Toxicol 17: abstract 529.	This report uses information from a base set of 438 notifications (1984-1997). One of the questions addressed was “how often does acute dermal toxicity contribute to overall classification?” The results showed that they contributed to for only about 1% of the chemicals tested. In only one instance the acute dermal study resulted in more severe classification than the oral study.
Thomas, H.D., Dewhurst, I.C., 2007. What Does a Dermal Acute Toxicity Study Add to the Information on a Plant Protection Pesticide? Abstracts Toxicology 231: 114 – 115.	The authors undertook an evaluation on the information on acute toxicity results and classification for 195 pesticides active substances evaluated within the EU since 1996 and 3111 formulated products approved by the Pesticides Safety Directorate (PSD). The results showed that 85% of the active substances and 92% of the formulated products merit no classification for dermal hazard. Only two active substances out of the 195 evaluated (1%) were more severely classified by dermal than oral route, one of which was also classified as ‘causes burns’. Only 9 out of 3111 products (0.3%) had a higher classification by dermal route than via other routes of exposure. It was concluded that the dermal acute toxicity study adds little if anything to the database on pesticide active substances. In all cases an equivalent indication of hazard, although often by a different route, was present.
Creton, S., Dewhurst, I.C., Earl, L.K., Gehen, S.C., Guest, R.L., Hotchkiss, J.A., Indans, I., Woolhiser, M.R., Billington, R., 2010. Acute toxicity testing of chemicals – Opportunities to avoid redundant testing and use alternative approaches. Critical Reviews in Toxicology 40(1): 50-83.	This paper reviews existing arguments for redundancy in acute toxicity testing of chemicals and chemical preparations, and the potential for use of alternative methodologies in the generation of acute toxicity data, giving consideration to the sometimes disparate regulatory approaches in place across various industry sectors. With regard to acute oral and dermal toxicity the authors reviewed 240 pesticide substances and demonstrated concordance of approaching 100% between substances not classified for acute toxicity by the oral or dermal routes; there was only one exception and this was a substance which demonstrated local corrosive effects in the dermal toxicity study. The authors also reviewed data for 438 industrial chemicals and report that only a single substance, whose identity was unknown, was classified for acute dermal toxicity but not for acute oral toxicity.
EPAA, Annual Report 2010	The document gives an overview of EPAA main activities during 2010. With regard to acute systemic toxicity a summary is provided on a workshop involving regulators, academics, poison centre staff and industry representatives that was held in Brussels on 16 September 2010. The focus was on two areas stemming from the paper “Cross-Sector Review of Drivers and Available 3Rs Approaches for Acute Systemic Toxicity Testing”, T. Seidle et al, in Toxicological Sciences: assessing 1) whether there are scientific drivers for acute toxicity testing and 2) whether the available concordance data set is adequate to support the deletion of the requirement for dermal testing from relevant legislation. Specific proposals were made for waivers that would deliver direct 3Rs benefits. The outcome of the workshop was communicated in a workshop report. Follow up steps included a wider consultation of the key recommendations with regulatory

	stakeholders and considerations of how the removal of the dermal route as part of default testing can be implemented.
Seidle, T., Robinson, S., Holmes, T., Creton, S., Prieto, P., Scheel, J., Chlebus, M., 2010. Cross-sector review of drivers and available 3Rs approaches for acute systemic toxicity testing. <i>Toxicological Sciences</i> 116(2): 382-396.	Based on the experience made within the pharmaceutical industry with acute oral toxicity studies, this review aimed to identify, a) regulatory and scientific drivers for acute toxicity testing in other industrial sectors, b) activities aimed at replacing, reducing, or refining the use of animals, and c) to make recommendations for future work in the area of acute toxicity testing. The work was carried out by the EPAA task force established to examine scientific and regulatory drivers for such testing and to promote the use of 3Rs approaches. Reference is also made to the results of an EPAA survey that included questions on the scientific and regulatory objectives of the studies, routes of administration, preferred test guideline, parameters examined, dose limit, and regulatory experience.
Andrew, D., Wright-Williams, S., 2011. Analysis of acute toxicity and irritation data submitted under the first REACH deadline. TSGE Poster presented at 2011 Autumn meeting of the British Toxicology Society.	This analysis over 100 substances registered under the first REACH deadline demonstrated that all of the substances not classified for acute oral toxicity were also not classified for acute dermal toxicity. Two substances were found to have more severe classification for acute dermal than acute oral toxicity. Both of them were highly toxic and corrosive inorganic substances and, therefore, the authors concluded that the dermal toxicity of both of these substances is likely to be a combination of severe local corrosivity and systemic effects.
Seidle, T., Prieto P, Bulgheroni A. 2011. Examining the Regulatory Value of Multi-route Mammalian Acute Systemic Toxicity Studies. <i>Alternatives to Animal Experimentation</i> 28(2): 95-102.	This paper describes the results of the analysis performed to assess oral-dermal and oral inhalation concordances among regulatory classifications for large data sets of chemicals (1569) and pesticide active ingredients (337) to determine the value of acute toxicity testing by more than one route. Overall the analysis provides further evidence that dermal acute systemic toxicity data almost never drive regulatory classification and labelling decisions in the chemicals, agrochemicals, or biocides sectors. With regard to industrial chemicals the oral-dermal classification concordance was 93.7% (1470/1569). The oral route lead to more toxic classification than the dermal route in 6.2% (98/1569) of cases and the dermal route lead to more toxic classification for only one substance out of the 1569 evaluated (0.06%). In addition, 1469 out of 1569 substances were not classified for acute toxicity (93%) and in all these cases the oral and dermal routes lead to no classification (none of the substances which were not classified for acute oral toxicity were classified for acute dermal toxicity). It was recommended to conduct an <i>in vitro</i> dermal absorption/penetration study (OECD, 2004) before carrying out a new dermal acute toxicity study to assess the likely magnitude and rate of dermal bioavailability. The differences seen in acute toxic responses following oral and inhalation exposure need further investigation.
EPAA Annual report 2012	The document gives an overview of EPAA main activities during 2012. With regard to acute systemic toxicity it provides a summary of the three Workstreams: 1) Reducing animal use in

	<p>dermal safety evaluation of substances, 2) How could C&L decisions in the agrochemical and chemical sectors be made if stand-alone acute toxicity testing were prohibited?, and 3) Minimising the requirements for inhalation testing.</p>
<p>Moore, N.P., Andrew, D.J., Bjerke, D.L., Creton, S., Dreher, D., Holmes, T., Prieto, P., Seidle, T., Rowan, T.G., 2013. Can acute dermal systemic toxicity tests be replaced with oral tests? A comparison of route-specific systemic toxicity and hazard classifications under the Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Regul. Toxicol. Pharmacol. 66(1): 30-37.</p>	<p>In this paper acute systemic toxicity data (LD₅₀ values) and hazard classifications derived in the rat following oral administration and dermal application have been analysed to examine whether or not orally-derived hazard classification or LD₅₀ values can be used to determine dermal hazard classification (335 substances analysed). The challenges associated with comparing route-specific classifications are discussed. The comparison of oral and dermal classifications resulted in 17% concordance, with 7% of substances classified less severely and 76% more severely if oral classifications were applied directly to the dermal route. In contrast, applying the oral LD₅₀ values within the dermal classification criteria to determine the dermal classification reduced the concordance to 15% and the relative 'under-classification' to 1%, but increased the relative 'over-classification' to 84%. The authors concluded that toxicity of chemicals is greater by the oral route than the dermal route and that oral acute systemic toxicity data can be used in place of equivalent dermal testing, with little or no concern for under-classification, by application of the dermal classification criteria. This approach does increase the potential for over-classification, which may be ameliorated by consideration of dermal penetration. Furthermore, testing a substance with an oral LD₅₀ value of >2000 mg/kg by the dermal route will not add to its hazard characterisation and should be dissuaded.</p>

7. Repeat dose systemic toxicity

Pilar Prieto & Andrew Worth

7.1 *The traditional animal test and its regulatory use*

Repeat dose toxicity studies in animals are performed to characterise the toxicological profile of a test substance in mammalian species following daily administration of the substance to be tested for a defined period of time up to the whole lifespan of the animals (i.e. sub-acute, sub-chronic and chronic exposures). There are different regulatory purposes, and the degree of uncertainty tolerated from the assessment varies accordingly; however the ultimate aim is to assess potential adverse effects on humans. These tests provide information on possible adverse effects on target organs, on dose-response relationships, and on the reversibility/irreversibility of the effects. In addition, they can be used to demonstrate delayed and accumulative effects. Information on a wide range of endpoints is provided, including histological changes in many organs and tissues, clinical signs, clinical chemistry and haematology. These *in vivo* tests generate the no-observed-adverse-effect level (NOAEL) or benchmark dose that is used as the point of departure in human risk assessment (ECHA, 2012). The study types are described in several OECD test guidelines [OECD, 1981a,b, 1996, 1998a,b,c, 2008, 2009a,b,c,d] and EU Test methods (EC, 2008, EU, 2014) including 28 day oral, dermal and inhalation studies in rodents, 90 day oral, dermal and inhalation studies in rodents, 90 day oral study in non-rodents, and chronic toxicity studies in rodents (OECD TGs 407-413, 452, EU test methods B.7 – B.9, B.26-B.30). The current revised OECD TG 407 allows the identification of certain endocrine mediated effects, as well as chemicals with neurotoxic potential and those that interfere with thyroid physiology. It may also warn on immunological effects. Other studies include combined repeat dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422) and the combined chronic toxicity/carcinogenicity studies (OECD TG 453, EU Test Method B.33).

The regulatory hazard/risk assessment procedures for industrial chemicals (European Commission, 2006), agrochemicals (European Commission, 2009a; European Commission, 2013), biocides (European Union, 2012), pharmaceuticals (EMA, 2010; ICH, 2010) and cosmetics (European Commission, 2009b; SCCS, 2012) require repeat dose toxicity testing to assess the potential risks of a substance.

7.2 *Mechanistic understanding of the endpoint*

The *in vivo* consequences of the repeat/chronic exposure to chemicals involve integrated processes at the molecular, cellular, organ and system levels. Overall, there is limited knowledge of the underlying mechanistic pathways and their interactions, despite several decades of research using *in vitro* and *in vivo* models. Many compounds are thought to induce their toxic effects by interfering with basic biochemical functions (e.g. interference with energy production, DNA function, receptor-mediated signalling pathways) or homeostatic mechanisms (e.g. perturbation of calcium homeostasis), resulting in functional impairments at the cell, tissue and organ levels. However, a quantitative description of these effects (dose-response relationships) is generally lacking.

Nevertheless, for some mechanisms of target organ systemic toxicity, a substantial body of knowledge already exists. For instance, due to its location and function in the organism, the liver is a sensitive target organ for several substances and considerable attention has been paid over the years to decipher the precise mechanisms underlying the induced adverse liver effects, such as necrosis, fibrosis, steatosis, phospholipidosis and cholestasis. Several recent review papers describe the molecular mechanisms of liver toxicity (Rusmann et al., 2009; Gomez-Lechon et al., 2010). In addition to the liver, other common target organs and systems such as the kidney, lung, central and peripheral nervous systems, cardiovascular system, and

immune system have been also the subject of several mechanistic investigations over the years (Prieto et al., 2006; Adler et al., 2011; Basketter et al., 2012).

Recently, in line with the mode of action/adverse outcome pathway (MoA/AOP) approach to toxicity prediction, work has started on the elucidation of the AOPs for some effects, such as fibrosis, steatosis and cholestasis (Landesmann et al., 2012; Vinken et al., 2013). The basic idea is that the knowledge of the underlying mechanisms simplified and represented in the form of AOPs, will enable the rational and flexible design of predictive systems. This approach is being explored by the SEURAT-1 consortium (SEURAT-1, 2011). A web-based platform (<http://moa-kb.jrc.ec.europa.eu/>) for sharing and developing AOPs is being developed by the JRC and EPA, in collaboration with the OECD and WHO IPCS.

7.3 Status of non-standard methods and Integrated Testing Strategies

7.3.1 *In vitro* methods

The availability and status of alternative methods for repeat dose toxicity has recently been reviewed in the context of the Cosmetics Directive 2013 marketing ban deadline (Adler et al., 2011). The expert group, convened by the JRC, concluded that although *in vitro* models are available for the most common target organs and systems such as the liver, kidney, lung, central nervous system, cardiovascular system and immune system (see Table 7.1 for summary of main approaches), none of them can be applied for quantitative risk assessment for repeat dose toxicity. However, for a limited number of target organs, *in vitro* models may be useful for mechanistic investigations and for hazard identification. In applying these models for hazard identification, it is important to understand whether chemically-induced perturbations observed *in vitro* (i.e. biochemical changes, up and down regulation of genes) are real biomarkers of toxicological effect, or simply adaptive responses (SEURAT-1, 2011; Blaauboer et al., 2012), and the potential differences between *in vitro* and *in vivo* responses. In this context, and following, a biomarker is defined a parameter that provides quantitative information that is mechanistically relevant to and predictive of an adverse effect (Boekelheide & Schuppe-Koistinen, 2012).

7.3.2 Multi-parametric approach

Several authors (Rusmann et al., 2009; Gomez-Lechon et al., 2010) have advocated the use of *in vitro* multi-parametric screening assays covering a wide spectrum of (mechanistically-based) key events, not only for mechanistic investigations but also for the prediction of systemic toxicities. These assays can include novel imaging technologies, which provide the possibility for continuous observation of critical cellular level events such as cell morphology, proliferation, migration, cell-cell interactions and colony formation. For example, multi-parametric assays for mechanistic and predictive studies have recently been proposed for drug-induced liver toxicity (hepatotoxicity) (Tolosa et al., 2012) and for mitochondrial dysfunction (Porceddu et al., 2012), a major mechanism of drug-induced liver injury.

7.3.3 Omics

Another major area of research and development concerns the wide range of “omics” techniques, which provide a means of understanding mechanisms of action at the molecular level, such as chemically-induced gene expression (Heijne et al., 2005; Jennings et al., 2013; Wilmes et al., 2011). While these technologies hold considerable promise, further development is needed before they can be applied in quantitative risk assessment for repeat dose toxicity. Stable and reproducible long-term culture systems need to be developed, along with optimised protocols. One of the reported limitations of using well-established *in vitro* primary hepatocytes in toxicogenomics is in fact the significant discrepancy between test compound-induced gene expression alterations in the liver *in vivo* and in hepatocytes *in vitro*. In this regard, Schug et al. (2013) have recently shown that *in vitro* and *in vivo* discrepancies

in the chemically-induced expression of three metabolising genes can be explained by the rapid decrease of the orally administered chemical *in vivo* (i.e. short half-life) in contrast to the slow decrease of the concentrations of the chemical in the culture medium. These pharmacokinetic differences between the two systems clarified why the genes were still deregulated after 24 h exposure *in vitro*, but not *in vivo*.

One of the challenges ahead is to study the interactions between genes, proteins and metabolites and to integrate the results. Some authors have already presented the benefits of an integrated *in vitro* omics approach in repeat dose toxicity assessment, combining transcriptomics, proteomics and metabolomics together with pharmacokinetics (e.g. Wilmes et al., 2013) and systems biology models (Hamon et al., 2014).

7.3.4 *In silico* and QSAR models

The availability of QSAR models for repeat dose toxicity endpoints is limited, as reviewed by Lapenna et al. (2010). Efforts have focussed on the use of molecular descriptors and statistically based modelling for the prediction of dose descriptors (e.g. prediction of the Lowest Observed Adverse Effect Level (LOAEL) or the Maximum Recommended Therapeutic Dose (MRTD) value), or of organ-related toxicities, including hepatotoxicity, nephrotoxicity, neurotoxicity and cardiotoxicity. For example, Myshkin et al. (2012) have developed models for hepatotoxicity and nephrotoxicity, as well as more precisely defined specific toxicity types for these organs (e.g. liver necrosis, kidney necrosis). The models for hepatotoxicity and nephrotoxicity were less accurate than those developed for more specific toxicity types within an organ. A more recent trend, in line with the MoA/AOP concept, has been to develop or use models for certain key events along a known toxicity pathway, for example binding to the LXR receptor (Spreafico et al, 2010), to be potentially used as a building block in a future AOP.

The most commonly used QSAR software tool for overall LOAEL prediction is probably TOPKAT (<http://www.accelrys.com>), which is a commercial product. Despite the lack of transparency in its predictions, several studies (e.g. Tilaoui et al, 2007) have claimed that TOPKAT gives reasonable predictions for a range of chemicals (including industrial chemicals). A more recently commercialised tool for LOAEL prediction is MolCode Toolboxes (<http://molcode.com/>). MRTD prediction can be carried out by using ADMET Predictor (<http://www.simulations-plus.com/>), MultiCase (<http://www.multicase.com/>) or the freely available software tool, Lazar (<http://lazar.in-silico.de>).

7.3.5 *Threshold of Toxicological Concern*

Another non-animal approach, which has gained increasing acceptance in the assessment of chemicals (especially non-intentionally added substances) in food (EFSA, 2012a) and cosmetics (SCHER/SCCS/SCENIHR, 2012), is based on the Threshold of Toxicological Concern (TTC) concept. The TTC concept is based on empirical evidence that for non-cancer effects, there are thresholds (exposure levels) below which toxicity does not occur, whereas for cancer effects, the likelihood of tumours is zero to very small at very low exposure levels. Thus, for chemicals of unknown toxicity, human exposure thresholds can be established below which there is a low probability of adverse effects on health. Accordingly, a range of human exposure thresholds (TTC values) have been developed for both cancer and non-cancer endpoints, on the basis of data from extensive toxicological testing in animals. For non-cancer effects (repeat dose toxicities), the most commonly accepted TTC values were proposed by Munro et al (1996) and are related to structural classes by the Cramer classification scheme (Lapenna et al, 2011). Specifically, Cramer classes I, II and III, respectively relating to low, medium and high concern for oral systemic toxicity, were associated exposure thresholds of 1800, 540 and 90 µg/person/day, respectively. More

recently, in order to address all types of populations, it has been considered that the thresholds should be expressed in $\mu\text{g}/\text{kg}$ body weight (bw) per day. Based on the (historical) assumption of a 60 kg adult, the corresponding thresholds for Cramer classes I, II and III are 30, 9, and 1.5 $\mu\text{g}/\text{kg}$ bw per day.

7.3.6 Challenges of the non-animal approaches

The difficulty in developing non-animal approaches for repeat dose toxicity is related to complexity of the underlying processes, which includes effects at different levels of biological organisation and time scales. *In vitro* methods have generally been developed to predict effects in specific target organs, without considering how the information generated can be used to predict the complex *in vivo* interactions between different cell types and tissue systems leading to toxicity.

Part of the solution lies in the use of Physiologically Based Biokinetic (PBBK) models. These are mathematical models which describe the flow of substances between the interconnected compartments (organs and tissues) of the whole organism, and which generate dose and time-dependent internal concentration profiles (e.g. internal concentrations at target organs), PBBK models rely on multiple input parameters, including substance-independent physiological parameters such as organ weights and blood flows, and substance-dependent parameters such as partitioning coefficients (e.g. blood-tissue partitioning) and metabolic fate (e.g. rate constants). PBBK models can be used to extrapolate biological effects between species, individuals, routes of exposure, and time scales (Coecke et al., 2013). They can also be used for *in vitro* to *in vivo* extrapolation. In this context, most studies have used the nominal (applied) *in vitro* concentration as the starting point for the extrapolation, but this can introduce inaccuracies due to the fate of the chemical in the *in vitro* system, such as binding to plastic, cell membranes and serum proteins, reducing bioavailability inside the cell (Broeders et al., 2013; Wilmes et al., 2013; Zaldivar et al., 2011). Therefore, recent efforts have focussed on modelling the *in vitro* system as well, and introducing a correction into the *in vitro* to *in vivo* extrapolation for the specific substance. An example is the virtual cell-based assay developed by the JRC (Zaldivar et al., 2011). A more complex solution is to integrate directly the dynamics of target tissue response, and use multi-scale modelling to translate the effects across multiple levels of biological organisation [from the cell to the whole body and timescales (Diaz et al, 2012)]. These developments are still at the research stage, and not yet ready for routine application in regulatory decision making.

7.4 Integrated testing strategies

Despite the limitations of individual alternative methods, it is widely considered that there are opportunities to avoid *in vivo* testing by optimal use of predictive methods such as QSARs, read-across, TTC and *in vitro* methods (Grindon et al., 2008; Adler et al., 2011). Several research initiatives around the world have taken up the challenge of combining and interpreting data on multiple targets generated by a variety of non-animal approaches.

In the USA, multiple initiatives are underway, including the Tox21 programme (<http://www.epa.gov/ncct/Tox21/>), the ToxCast screening programme (<http://www.epa.gov/ncct/toxcast/>) and work led by the Hamner Institute for Chemical Safety Sciences (<http://www.thehamner.org/institutes-centers/institute-for-chemical-safety-sciences/>). These initiatives are developing new assays to characterise major “toxicity pathways,” dose response models to quantify when perturbations in toxicity pathways result in adverse events in cells and tissues, and *in vitro* to *in vivo* extrapolation tools to predict the level and duration of exposure that is likely to result in adverse events in humans.

In Europe, the FP7 project Predict-IV (www.predict-iv.toxi.uni-wuerzburg.de/) aimed to combine *in vitro* biological effects (determined by transcriptomics, proteomics and metabolomics, and high content imaging) with pharmacokinetics and modelling (*in vitro* to *in vivo* extrapolation) in order to develop testing strategies based on mechanistic information that could improve pre-clinical safety testing of pharmaceuticals. The emphasis of SEURAT-1 research strategy is on the identification and elucidation of MoA related to repeat dose systemic toxicity in humans and the development of experimental and computational models that capture the related key events and toxicity pathways (<http://www.seurat-1.eu/>; SEURAT-1, 2012, 2013).

7.5 Conclusions

Information on the repeat dose effects of chemicals is required for human health hazard assessments under REACH and the Biocides Products Regulation, and can be used for classification and labelling under CLP.

The development of reliable and useful non-animal approaches for repeat dose toxicity has been hampered by an incomplete understanding of the complex network of biological processes involved, as well as the technical challenge of simulating the most relevant processes *in vitro* or *in silico*. These processes include both the biokinetics of internal exposure in whole organisms and the dynamics of effects in target organs. Until recently, the need to explicitly account for *in vivo* biokinetics (by using PBBK models) has largely been ignored, with efforts focusing on the development of methods predicting effects in target organs. Due to the potentially large number of targets and mechanisms, it is however unlikely that any single *in vitro* test or QSAR model will be capable of making reliable predictions for all chemicals of interest.

A wide range of *in vitro* models are available, but none generates reliable quantitative information suitable for use in hazard assessment. However, for a limited number of target organ toxicities, *in vitro* models may be useful for hazard identification and to obtain mechanistic information. The application of multi-parametric, high-throughput screening (HTS), high content imaging (HCI) and omics technologies to *in vitro* models is a broad and promising area of research, although at present, these approaches are not sufficiently well developed or practicable for routine use in quantitative hazard assessment. These approaches do however provide a powerful means of hypothesis building and informing the development of more targeted and simpler assays. For example, omics techniques can be used to explore the time-dependencies of differential gene expression, and to establish which genes and time point(s) are most relevant for predicting a given adverse outcome.

A limited number of QSAR models are available. QSARs for LOAEL prediction are of questionable relevance, since they do not conform with the OECD validation principle of predicting a “defined endpoint”; there is no convincing mechanistic basis for the statistical association between a set of molecular descriptors and the potencies of lead effects arising in different organs. More credible QSARs, more in line with 21st Century thinking, need to be developed for the prediction of specific key events in AOPs, which provide a rational basis for integrating their predictions with experimental data generated by *in vitro* methods.

In view of the limited availability of QSARs and PBBK models, the well-established chemical category/read-across approach is of key importance. Particularly interesting in this regard is the possibility of filling data gaps by using read-across based on both structural and biological descriptors, sometimes referred to as “chemical biological” read-across (CBRA; Low et al, 2013). In this context, any toxicologically relevant read-out (including functional parameters and omics measurements) from an *in vitro* system could be a useful biological descriptor. This topic merits further investigation.

PBBK models have been developed for a limited number of chemicals, and are increasingly being integrated with toxicological data from *in vitro* studies. However, further research is needed to expand the repertoire of such models and integrate them with relevant *in vitro* data to make quantitative predictions of toxicological effects by extrapolating from *in vitro* to *in vivo* points of departure. It will also be important to implement these models in user-friendly platforms to make them more accessible for routine use in risk assessment.

In the assessment of industrial chemicals and biocides, the TTC approach could be used alongside QSAR in the assessment of degradates and metabolites for which experimental toxicity data are neither available nor explicitly required. To explore this further, a good starting point would be the guidance developed by EFSA for the assessment of pesticide metabolites (EFSA 2012b). Within current research projects (e.g. the FP7 project COSMOS), efforts are underway to further develop the TTC approach.

For quantitative risk assessment, an integrated approach based on an understanding of MoA and perturbation of biological pathways, as well as a better knowledge of *in vitro* and *in vivo* kinetics, and target organ/tissue dose-response relationships is needed (Adler et al.; 2011; SEURAT-1, 2011). A possible concern for regulatory implementation is that in moving towards a new toxicological paradigm will entail replacing a checklist of standard animal studies with a possibly more extensive checklist of more *ad hoc in vitro* and modelling studies. However, the vision and hope is that by employing the MoA/AOP framework, it should be possible to identify a limited number of key events that act as common “nodes” in multiple toxicity pathways, with the practical benefit that a limited set of (*in vitro* and *in silico*) models based on these key events will provide a sufficient toolkit for quantitative hazard assessment based entirely on non-animal approaches. A suite of AOPs relevant to repeat dose toxicity, including some of those from the above-mentioned research initiative, will be evaluated and published by the OECD.

7.6 References

- Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485
- Basketter DA, Clewell H, Kimber I, Rossi A, Blaauboer B, Burrier R, Daneshian M, Eskes C, Goldberg A, Hasiwa N, Hoffmann S, Jaworska J, Knudsen TB, Landsiedel R, Leist M, Locke P, Maxwell G, McKim J, McVey EA, Ouédraogo G, Patlewicz G, Pelkonen O, Roggen E, Rovida C, Ruhdel I, Schwarz M, Schepky A, Schoeters G, Skinner N, Trentz K, Turner M, Vanparys P, Yager J, Zurlo J & Hartung T (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report. *ALTEX* 29(1): 3-91.
- Blaauboer BJ, Boekelheide K, Clewell HJ, Daneshian M, Dingemans MM, Goldberg AM, Heneweer M, Jaworska J, Kramer NI, Leist M, Seibert H, Testai E, Vandebriel RJ, Yager JD & Zurlo J (2012). The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. *ALTEX* 29(4): 411-25.
- Boekelheide K & Schuppe-Koistinen I (2012). SOT/EUROTOX debate: biomarkers from blood and urine will replace traditional histopathological evaluation to determine adverse responses. *Toxicological Sciences* 129: 249–255.
- Broeders JJW, Blaauboer BJ & Hermens JLM (2013). In vitro biokinetics of chlorpromazine and the influence of different dose metrics on effect concentrations for cytotoxicity in Balb/c 3T3, Caco-2 and HepaRG cell cultures. *Toxicology in Vitro* 27: 1057–1064.
- Coecke S, Pelkonen O, Leite SB, Bernauer U, Bessems JG, Bois FY, Gundert-Remy U, Loizou G, Testai E & Zaldivar JM (2013). Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicol. In Vitro* 27(5): 1570-1577.
- Diaz Ochoa JG, Bucher J, Péry AR, Zaldivar Comenges JM, Niklas J & Mauch K (2012). A multi-scale modeling framework for individualized, spatiotemporal prediction of drug effects and toxicological risk. *Frontiers in Pharmacology* 3: 204.
- European Chemicals Agency (2012). Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific guidance. Available from: http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf
- European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official Journal of the European Union* L396, 1-849.
- European Commission (2009a). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. *Official Journal of the European Union* L309, 1-47.
- European Commission (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. *Official Journal of the European Union* L342, 59-209.

European Commission (2013). Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.

EFSA (2012a). Scientific Opinion on exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC) EFSA Journal 2012; 10(7):2750. European Food Safety Authority. Available from: <http://www.efsa.europa.eu/en/efsajournal/doc/2750.pdf>

EFSA (2012b) Scientific Opinion on evaluation of the toxicological relevance of pesticide metabolites for dietary risk assessment EFSA Journal 2012; 10(07):2799. European Food Safety Authority. Available from: <http://www.efsa.europa.eu/en/efsajournal/doc/2799.pdf>

European Medicines Agency (2010). Guideline on repeated dose toxicity. Committee for Human Medicinal Products. Reference number CPMP/SWP/1042/99 Rev

European Union (2008). Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L142.

European Union (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.

European Union (2014). Commission Regulation (EU) No 260/2014 of 24 January 2014 amending, for the purpose of its adaptation to technical progress, Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L 81, 1-253.

Gómez-Lechón MJ, Lahoz A, Gombau L, Castell JV & Donato MT (2010). In vitro evaluation of potential hepatotoxicity induced by drugs. *Curr Pharm Des.* 16(17): 1963-1977.

Grindon C, Combes R, Cronin MTD, Roberts DW & Garrod JF (2008). An Integrated Decision-tree Testing Strategy for Repeat Dose Toxicity with Respect to the Requirements of the EU REACH Legislation. *ATLA* 36: 93–101

Hamon J, Jennings P & Bois FY (2014). Systems biology modeling of omics data: effect of cyclosporine a on the Nrf2 pathway in human renal cells. *BMC Syst Biol.* 8(1): 76.

Heijne WH, Kienhuis S, van Ommen B, Stierum RH & Groten JP (2005). Systems toxicology: applications of toxicogenomics, transcriptomics, proteomics and metabolomics in toxicology. *Expert Rev Proteomics* 2(5): 767-80.

ICH (2010). International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals M3(R2). Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M3_R2/Step4/M3_R2_Guideline.pdf

Jennings P, Limonciel A, Felice L & Leonard MO (2013). An overview of transcriptional regulation in response to toxicological insult. *Arch. Toxicol.* 87: 49-72.

Landesmann B, Goumenou M, Munn S & Whelan M (2012). Description of prototype modes-of-action related to repeated dose toxicity. JRC Scientific and Policy Report EUR 25631 EN.

Publications Office of the European Union, Luxembourg. Available from: <http://publications.jrc.ec.europa.eu/repository/>

Lapenna S, Fuart-Gatnik M & Worth A (2010). Review of QSAR models and software tools for predicting acute and chronic systemic toxicity. JRC Scientific and Technical report EUR 24639 EN. Publications Office of the European Union, Luxembourg. Available from: <http://publications.jrc.ec.europa.eu/repository/>

Lapenna S & Worth A (2011). Analysis of the Cramer classification scheme for oral systemic toxicity - implications for its implementation in Toxtree. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: <http://publications.jrc.ec.europa.eu/repository/>

Low Y, Sedykh A, Fourches D, Golbraikh A, Whelan M, Rusyn I & Tropsha A (2013). Integrative chemical-biological read-across approach for chemical hazard classification. *Chemical Research in Toxicology* 26: 1199-1208.

Myshkin E, Brennan R, Khasanova T, Sitnik T, Serebriyskaya T, Litvinova E, Guryanov A, Nikolsky Y, Nikolskaya T & Bureeva S (2012). Prediction of Organ Toxicity Endpoints by QSAR Modeling Based on Precise Chemical-Histopathology Annotations. *Chemical Biology & Drug Design* 80(3): 406-416.

Munro IC, Ford RA, Kennepohl E & Sprenger JG (1996). Correlation of structural class with No-Observed-Effect Levels: A proposal for establishing a threshold of concern. *Food Chem Toxicol* 34: 829-867.

OECD (1981a). Test No. 410: Repeated Dose Dermal Toxicity: 21/28-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070745-en

OECD (1981b). Test No. 411: Subchronic Dermal Toxicity: 90-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070769-en

OECD (1996). Test No. 422: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070981-en

OECD (1998a). Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070707-en.

OECD (1998b). Test No. 409: Repeated Dose 90-Day Oral Toxicity Study in Non-Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070721-en.

OECD (2008). Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070684-en

OECD (2009a). Test No. 412: Subacute Inhalation Toxicity: 28-Day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070783-en

OECD (2009b). Test No. 413: Subchronic Inhalation Toxicity: 90-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070806-en

OECD (2009c). Test No. 452: Chronic Toxicity Studies, OECD Guidelines for the Testing of Chemicals. Section 4, OECD Publishing. doi: 10.1787/9789264071209-en.

OECD (2009d). Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071223-en

Prieto P, Baird AW, Blaauboer BJ, Castell Ripoll JV, Corvi R, Dekant W, Dietl P, Gennari A, Gribaldo L, Griffin JL, Hartung T, Heindel JJ, Hoet P, Jennings P, Marocchio L, Noraberg J, Pazos P, Westmoreland C, Wolf A, Wright J & Pfaller W (2006). The assessment of repeated dose toxicity in vitro: a proposed approach. The report and recommendations of ECVAM workshop 56. ATLA 34: 315-341.

Porceddu M, Buron N, Roussel C, Labbe G, Fromenty B & Borgne-Sanchez A (2012). Prediction of Liver Injury Induced by Chemicals in Human with a Multiparametric Assay on Isolated Mouse Liver Mitochondria. Toxicological Sciences 129: 332-345.

Russmann S., Kullak-Ublick GA & Grattagliano I (2009). Current Concepts of Mechanisms in Drug-Induced Hepatotoxicity. Current Medicinal Chemistry 16: 3041-3053

SCCS (2012). The SCCS's notes of guidance for the testing of cosmetic ingredients and their safety evaluation. Scientific Committee on Consumer Safety. Available from: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_006.pdf

SCHER/SCCS/SCENIHR (2012). Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products. Available from: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_092.pdf

SEURAT-1 (2011). Towards the Replacement of in vivo repeated Dose Systemic Toxicity Testing. Volume 1. Available from: http://www.seurat-1.eu/media/download_gallery/SEURAT-1_Annual%20Report_Vol%201_Sept2011_HR.pdf

SEURAT-1 (2012). Towards the Replacement of in vivo repeated Dose Systemic Toxicity Testing. Volume 2. Available from: [http://www.seurat-1.eu/media/download_gallery/SEURAT%201v2\(BD\).pdf](http://www.seurat-1.eu/media/download_gallery/SEURAT%201v2(BD).pdf)

SEURAT-1 (2013). Towards the Replacement of in vivo repeated Dose Systemic Toxicity Testing. Volume 3. Available from: http://www.seurat-1.eu/media/download_gallery/SEURAT%201-v3%28bd.pdf

Schug M, Stöber R, Heise T, Mielke H, Gundert-Remy U, Godoy P, Reif R, Blaszkewicz M, Ellinger-Ziegelbauer H, Ahr HJ, Selinski S, Günther G, Marchan R, Blaszkewicz M, Sachinidis A, Nüssler A, Oberemm A & Hengstler JG (2013). Pharmacokinetics explain in vivo/in vitro discrepancies of carcinogen-induced gene expression alterations in rat liver and cultivated hepatocytes. Archives of Toxicology 87: 337-345.

Spreatico M, Smiesko M, Peristera O, Rossato G & Vedani A (2010). Probing Small-Molecule Binding to the Liver-X Receptor: A Mixed-Model QSAR Study. Molecular Informatics 29: 27-36.

Tilaoui L, Schilter B, Tran L-A, Mazzatorta P & Grigorov M (2007). Integrated Computational Methods for Prediction of the Lowest Observable Adverse Effect Level of Food-Borne Molecules. QSAR Comb. Sci.: 26: 102-108.

Tolosa, L., Pinto, S., Donato, M.T., Lahoz, A., Castell, J.V., O'Connor, J.E., & Gómez-Lechón, M.J., (2012). Development of a Multiparametric Cell-based Protocol to Screen and Classify the Hepatotoxicity Potential of Drugs. Toxicological Sciences 127(1): 187-198.

Vinken M, Landesmann B, Goumenou M, Vinken S, Shah I, Jaeschke H, Willett C, Whelan M & Rogiers V (2013). Development of an adverse outcome pathway from drug-mediated bile salt export pump inhibition to cholestatic liver injury. Toxicol Sci. 136(1): 97-106.

Wilmes A, Crean D, Aydin S, Pfaller W, Jennings P & Leonard MO (2011). Identification and dissection of the Nrf2 mediated oxidative stress pathway in human renal proximal tubule toxicity. *Toxicology in Vitro* 25: 613-622.

Wilmes A, Limonciel A, Aschauer L, Moenks K, Bielow C, Leonard MO, Hamon J, Carpi D, Ruzek S, Handler A, Schmal O, Herrgen K, Bellwon P, Burek C, Truisi GL, Hewitt P, Di Consiglio E, Testai E, Blaauboer BJ, Guillou C, Huber CG, Lukas A, Walter P, Mueller SO, Bois FY, Dekant W & Jennings P (2013). Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. *J Proteomics* 79: 180-194.

Zaldivar Comenges JM, Menecozzi M, Macko P, Rodriguez R, Bouhifd M & Baraibar Fentanes J (2011). A Biology-Based Dynamic Approach for the Modelling of Toxicity in Cell Assays: Part II: Models for Cell Population Growth and Toxicity. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: <http://publications.jrc.ec.europa.eu/repository/>

Table 7.1. *In vitro* methods for common target organs in repeat dose toxicity

<i>In vitro</i> tests available	Part of mechanism covered	Area(s) of application	Status	References
<p><i>Liver-derived in vitro systems:</i></p> <ul style="list-style-type: none"> • Liver slices • Immortalised liver based cell lines • Primary hepatocytes (suspensions, monolayer cultures, sandwich cultures) • 3D tissue constructs • Bioartificial livers • Co-cultures of various cell types with hepatocytes • Stem cell models 	<p>Cytotoxicity (including necrosis and apoptosis), cholestasis, steatosis, phospholipidosis, fibrosis</p>	<p>Hazard identification</p>	<p>R&D</p>	<p>Adler et al., 2011 Soldatow et al., 2013</p>
<p><i>Kidney-derived in vitro systems:</i></p> <ul style="list-style-type: none"> • Primary cell cultures • Renal cortical slices • Not transformed renal epithelial cell lines (from several species) • Human proximal tubule epithelial cell lines • 3D tissue constructs • Liver-kidney co-culture model in micro-fluid chip • Human Proximal Tubule-on-a-chip • Renal epithelial like cells derived from stem cells 	<p>Nephrotoxicity</p>	<p>Hazard identification</p>	<p>R&D</p>	<p>Adler et al., 2011 Choucha-Snouber et al., 2012 DesRochers et al., 2013 Ginai et al., 2013</p>

<i>In vitro</i> tests available	Part of mechanism covered	Area(s) of application	Status	References
<p><i>Cardiovascular system derived in vitro systems:</i></p> <ul style="list-style-type: none"> • Murine and human embryonic stem cell-derived cardiomyocytes • Primary cardiomyocytes from a variety of animal species • Cell lines • Engineered cardiac tissue 	<p>Cytotoxicity, alterations in cardiac action potential, effects on ion channels (primarily hERG), arrhythmia</p>	<p>Screening/hazard identification</p>	<p>R&D</p> <p>hERG assays used in pharmaceutical development</p>	<p>Adler et al., 2011 Sukardi et al, 2011 Abassi et al., 2012</p>
<p><i>Central and peripheral nervous system derived in vitro systems:</i></p> <ul style="list-style-type: none"> • Primary cultures • Re-aggregate cultures (or explant/slice culture) • Neuroblastoma and glioma cell lines • Human and rodent neural stem cell lines • Non-mammalian models (e.g. zebra fish, medaka or <i>C. elegans</i>) 	<p>Neurophysiology testing Neural electrical communication Neurotransmission Enzymatic activity Cell morphology Cell differentiation General biochemistry</p>	<p>Screening/prioritisation Hazard identification</p>	<p>R&D</p>	<p>Adler et al., 2011 Sukardi et al, 2011</p>
<p><i>Respiratory tract derived in vitro systems:</i></p> <ul style="list-style-type: none"> • Primary cultures from human and rodent tissue • Lung slices • Cell lines mimicking different cell types of the respiratory tract • Co-cultures models (e.g. epithelial and dendritic cells or macrophages) • 3D in vitro human models 	<p>Allergic asthma Cystic fibrosis Inflammation Remodeling Mucus secretion Ciliary clearance Active ion transport Irritation Adaptive immunity</p>	<p>Hazard identification</p>	<p>R&D</p>	<p>Adler et al., 2011 Sauer et al., 2013</p>

<i>In vitro</i> tests available	Part of mechanism covered	Area(s) of application	Status	References
<i>Immune system derived in vitro systems:</i> <ul style="list-style-type: none"> • Human whole blood • Human peripheral blood cells • Human, mouse and rat bone marrow progenitors such as granulocyte-macrophage and megakaryocyte progenitors • Haematopoietic stem cells 	Immunotoxicity Myelotoxicity	Hazard identification	R&D Scientifically validated for evaluating myelotoxicity (human and murine CFU-GM)	Adler et al., 2011

Table references

Abassi YA, Xi B, Li N, Ouyang W, Seiler A, Watzele M, Kettenhofen R, Bohlen H, Ehlich A, Kolossov E, Wang X, Xu X. (2012). Dynamic monitoring of beating periodicity of stem cell-derived cardiomyocytes as a predictive tool for preclinical safety assessment. *British Journal of Pharmacology* 165(5), 1424-41.

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485

Choucha-Snouber, L., Aninat, C., Grsicom, L., Madalinski, G., Brochot, C., Poleni, P.E., Razan, F., Guillouzo, C.G., Legallais, C., Corlu, A., & Leclerc, E. (2013). Investigation of ifosfamide nephrotoxicity induced in a liver-kidney co-culture biochip. *Biotechnology & Bioengineering* 110(2), 597-608.

DesRochers, T.M., Suter, L., Roth, A., & Kaplan, D.L. (2013). Bioengineered 3D human kidney tissue, a platform for the determination of nephrotoxicity. *PLoS One*. doi: 10.1371/journal.pone.0059219.

Ginai, M., Elsby, R., Hewitt, C.J., Surry, D., Fenner, K., & Coopman, K. (2013). The use of bioreactors as in vitro models in pharmaceutical research. *Drug Discovery Today*. <http://dx.doi.org/10.1016/j.drudis.2013.05.016>

Sauer, U.G., Vogel, S., Hess, A., Kollé, S.N., Ma-Hock, L., van Ravenzwaay, B., Landsiedel, R. (2013). In vivo-in vitro comparison of acute respiratory tract toxicity using human 3D airway epithelial models and human A549 and murine 3T3 monolayer cell systems. *Toxicology In Vitro* 27(1), 174-90.

Soldatow, V.Y., LeCluyse, E.L., Griffiths, L.G., & Rusyn, I. (2013). In vitro models for liver toxicity testing. *Toxicology Research* 2, 23-39.

Sukardi, H., Chng, H.T., Chan, E.C., Gong, Z., Lam, S.H. (2011). Zebrafish for drug toxicity screening: bridging the in vitro cell-based models and in vivo mammalian models. *Expert Opinion on Drug Metabolism & Toxicology* 7(5), 579-89.

Chapter 7. Table of References with Notes

Traditional animal tests – OECD TGs/EU test methods	
OECD (1996), Test No. 422: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070981-en	This method is designed for use with the rat and oral administration is assumed. It provides information on the possible health hazards likely to arise from repeated exposure over a relatively limited period of time. The method comprises the basic repeated dose toxicity study that may be used for chemicals on which a 90-day study is not warranted (e.g. when the production volume does not exceed certain limits) or as a preliminary study to a long-term study. It further comprises a reproduction/developmental toxicity screening test and also places emphasis on neurological effects as a specific endpoint.
OECD (1981a) Test No. 410: Repeated Dose Dermal Toxicity: 21/28-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070745-en	This method provides information on possible health hazards likely to arise from repeated exposures by the dermal route over a period of 21 or 28 days. The test substance is applied daily to the skin. The adult rat, rabbit or guinea pig may be used. Both sexes should be used.
EU test method B.9. Repeated dose (28 days) toxicity (dermal).	
OECD (1981b) Test No. 411: Subchronic Dermal Toxicity: 90-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070769-en	This method is based on the repeated application by the dermal route of the substance of interest during 90 days (several hours daily). It is intended for use with the adult rat, rabbit or guinea pig. Both sexes should be used.
EU test method B.28. Sub-chronic dermal toxicity study 90 day repeated dermal dose study using rodent species	
OECD (1998b) Test No. 409: Repeated Dose 90-Day Oral Toxicity Study in Non-Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070721-en.	This method is based on the repeated oral administration of the substance during 90 days (one dose level daily). It is intended primarily for use with non-rodent species (the dog and, in particular, the beagle is frequently used). Both sexes should be used.
EU test method B.27. Sub-chronic oral toxicity test repeated dose 90-day oral toxicity study in non-rodents	
OECD (1998a) Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070707-en	This method is based on the repeated oral administration of the substance of interest over 90 days (one dose level daily). It is intended primarily for use with rodents (rat preferably). Both sexes should be used.
EU test method B.26. Sub-chronic oral toxicity test repeated dose 90 – day oral toxicity study in rodents	
OECD (2008) Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents. OECD Guidelines for the Testing of Chemicals, Section 4, OECD	This method is based on the repeated administration of the substance of interest via the oral route during 28 days (one dose level daily). It is intended primarily

<p>Publishing. doi: 10.1787/9789264070684-en</p> <p>EU test method B.7. Repeated dose 28-day oral toxicity study in rodents</p>	<p>for use with rodents (rat preferably) and both sexes should be used. This method is intended to investigate effects on a very broad variety of potential targets of toxicity, including effects on the nervous, immune and endocrine systems. Regarding these particular endpoints, it should identify chemicals with neurotoxic potential, which may warrant further in-depth investigation of this aspect, and chemicals that interfere with thyroid physiology. It may also provide data on chemicals that affect the male and/or female reproductive organs in young adult animals and may give an indication of immunological effects.</p>
<p>OECD (2009c) Test No. 452: Chronic Toxicity Studies, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071209-en</p> <p>EU test method B.30. Chronic toxicity studies</p>	<p>The aim of this method is to characterise the profile of a substance in a mammalian species (primarily rodents) following prolonged and repeated exposure over 12 months. Frequency of exposure normally is daily, but may vary according to the route chosen (oral, dermal or inhalation) and should be adjusted according to the toxicokinetic profile of the test substance. Both sexes should be used.</p>
<p>OECD (2009d) Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071223-en</p> <p>EU test method B.33. Combined chronic toxicity/carcinogenicity studies</p>	<p>The objective of a combined chronic toxicity/carcinogenicity study is to identify carcinogenic and the majority of chronic effects, and to determine dose-response relationships following prolonged and repeated exposure. The rat is typically used for this study and focuses on the oral route of administration.</p>
<p>OECD (2009a) Test No. 412: Subacute Inhalation Toxicity: 28-Day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070783-en</p> <p>EU test method B.8. Subacute inhalation toxicity: 28-day study</p>	<p>This method aims to fully characterise the test substance toxicity by the inhalation route following repeated exposure during 28 days, and to provide data for quantitative inhalation risk assessments.</p>
<p>OECD (2009b) Test No. 413: Subchronic Inhalation Toxicity: 90-day Study. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264070806-en</p> <p>EU test method B. 29. Subchronic inhalation toxicity: 90-day study</p>	<p>This method aims to fully characterise the test substance toxicity by the inhalation route following repeated exposure during 90 days, and to provide robust data for quantitative inhalation risk assessments.</p>
<p>Regulatory requirements across sectors</p>	
<p>European Commission, 2006. Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive</p>	<p><i>Chemicals:</i> Information requirements depend on the production volume of the chemical concerned:</p> <ul style="list-style-type: none"> • Annex VIII (10-100 tpa): Short-term repeated dose toxicity study (28 days), one species, male and female, most appropriate route of administration, having regard to the likely route of human exposure. • Annex IX - X (100->1000 tpa): In addition to the 28-day repeated dose

76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Communities L396, 1-849.	toxicity study, a sub-chronic toxicity study (90 days) should be carried out in rodents, male and female, by the most appropriate route of administration, having regard to the likely route of human exposure.
European Commission, 2008. Council Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) AND European Union, 2014. Commission Regulation (EU) No 260/2014 of 24 January 2014 amending, for the purpose of its adaptation to technical progress, Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L 81, 1-253	The Regulation of May 2008 and its subsequent amendments, define tests testing of chemicals for the REACH Regulation. They are based on the OECD Guidelines for the Testing of Chemicals. The January 2014 adaptation contains, among others, nine methods for the determination of toxicity and other health effects including four inhalation toxicity test methods, which include an update of three methods and one new method to reduce the number of animals used and to improve assessment of effects, an update of the repeat dose 28-day oral toxicity test method to include parameters for assessment of endocrine activity, an update of the toxicokinetics test method relevant for the design and understanding of toxicological studies and an update of chronic, carcinogenicity and combined chronic and carcinogenicity test methods.
European Commission, 2009a. Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Communities L342, 59-209.	<i>Cosmetics</i> : This Regulation establishes rules to be complied with by any cosmetic product made available on the market, in order to ensure the functioning of the internal market and a high level of protection of human health. Article 18 (Animal testing)
European Commission, 2009b. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-47. AND European Commission, 2013. Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.	<i>Active substances</i> : Where available, oral 28-day studies shall be reported. Oral 90-day studies, usually the rat, and non-rodents (90-day toxicity study in dogs), shall always be reported. For human risk assessment additional dermal studies shall be considered on a case by case basis, unless the active substance is a severe irritant. For volatile active substances (vapour pressure >10 ⁻² Pascal) expert shall be required to decide whether the short term studies have to be performed by inhalation exposure. A long-term oral toxicity study of the active substance shall be conducted using rat as test species; where possible this study shall be combined with a long-term carcinogenicity study (two years).
European Medicines Agency, 2010. Guideline on repeated dose toxicity. Committee for Human Medicinal Products. Reference number CPMP/SWP/1042/99 Rev	This Guideline concerns the conduct of repeated dose toxicity studies of active substances intended for human use. It includes herbal medicinal products while for biotechnology-derived compounds, vaccines and anticancer medicinal products, specific guidance is available. In general, repeated dose toxicity studies shall be carried out in two species of mammals, one of which must be a non-rodent. The dose regimen, duration and route of administration should be chosen based on the intended clinical use in humans.

<p>ICH - International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use, 2010. Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals M3(R2).</p>	<p>The purpose of this document is to recommend international standards for, and promote harmonisation of, the nonclinical safety studies recommended to support human clinical trials of a given scope and duration as well as marketing authorization for pharmaceuticals. The recommended duration of the repeated-dose toxicity studies is usually related to the duration, therapeutic indication and scope of the proposed clinical trial. Repeated-dose toxicity studies in two species (one non-rodent) for a minimum duration of 2 weeks would generally support any clinical development trial up to 2 weeks in duration. Clinical trials of longer duration should be supported by repeated-dose toxicity studies of at least equivalent duration. Six month rodent and 9 month no-rodent studies generally support dosing for longer than 6 months in clinical trials.</p> <p>Safety pharmacology and pharmacodynamic (PD) studies are defined in ICH S7A. The core battery of safety pharmacology studies includes the assessment of effects on cardiovascular, central nervous and respiratory systems, and should generally be conducted before human exposure, in accordance with ICH S7A and S7B. When warranted, supplemental and follow-up safety pharmacology studies can be conducted during later clinical development. Consideration should be given to inclusion of any <i>in vivo</i> evaluations as additions to general toxicity studies, to the extent feasible, in order to reduce animal use. In addition, primary PD studies (<i>in vivo</i> and/or <i>in vitro</i>) are intended to investigate the mode of action and/or effects of a substance in relation to its desired therapeutic target. Such studies are generally conducted during the discovery phase of pharmaceutical development and as such, are not generally conducted in accordance with Good Laboratory Practice. These studies can contribute to dose selection for both nonclinical and clinical studies</p>
<p>European Chemicals Agency, 2012. Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific guidance. Available from: http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf</p>	<p><i>Chemicals:</i> This document describes the information requirements under REACH with regard to substance properties, exposure, uses and risk management measures, and the chemical safety assessment. For repeated dose toxicity, all available information relevant for the endpoint needs to be evaluated and classification considered at each tonnage level. The standard information requirements on repeated dose toxicity are specified in REACH Annexes VII-X.</p>
<p>European Union, 2012. Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.</p>	<p><i>Biocides:</i> In general, only one route of administration is necessary and the oral route is the preferred route. However, in some cases it may be necessary to evaluate more than one route of exposure. Short-term repeated dose toxicity study (28 days) and for the sub-chronic repeated dose toxicity study (90 days) the preferred species is rat. The duration of the long-term repeated dose toxicity is equal or higher to 12 months</p>

	Further repeat dose studies including testing on a second species (non-rodent), studies of longer duration or through a different route of administration shall be undertaken if certain conditions are met.
SCCS: Scientific Committee on Consumer Safety (2012): The SCCS's notes of guidance for the testing of cosmetic ingredients and their safety evaluation. Available from: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_006.pdf .	The 28-day and 90-day oral toxicity tests in rodents are the most commonly used repeated dose toxicity tests. Preferably studies of 90 days or more should be used in safety assessments. If studies of only 28 days duration are available, an additional uncertainty factor can be used in the calculation of the margin of safety. The inhalation route is only rarely used due to the complex study design, as well as to the lack of relevance of this route of repeated exposure for the majority of cosmetic products. In a number of cases, dermal repeated dose toxicity studies are present among the submitted data (e.g. UV-filter).
Current status and opportunities for the development of alternative approaches for systemic toxicity testing	
Heijne, W.H., Kienhuis, S., van Ommen, B., Stierum, R.H., Groten, J.P., 2005. Systems toxicology: applications of toxicogenomics, transcriptomics, proteomics and metabolomics in toxicology. <i>Expert Rev Proteomics</i> . 2(5), 767-80.	According to this review: <ul style="list-style-type: none"> • Toxicogenomics can facilitate the identification and characterisation of toxicity. • Toxicogenomics methods may reduce uncertainties by providing detailed insight into molecular mechanisms underlying the toxic endpoint. • Application of toxicogenomics to <i>in vitro</i> studies has the potential to reduce, refine and replace animal experiments. • Functional genomic methods will only be fully deployed when these methods are integrated with conventional techniques to allow an iterative process of hypothesis generation and confirmation. • Transcriptomics and metabolomics are sensitive and becoming increasingly reproducible, while proteomics methods may become so after further developments. • The current challenge is to study interactions between genes, proteins and metabolites, and to integrate the results of the different genomics methods using a systems toxicology concept.
Prieto, P, Baird, A W, Blaauboer, B J, et al., 2006. The assessment of repeated dose toxicity in vitro: a proposed approach. The report and recommendations of ECVAM workshop 56.- <i>ATLA</i> , 34, 315-341.	This report summarises the outcome of a workshop during which alternative approaches to <i>in vivo</i> repeated dose toxicity testing were discussed and evaluated by a selected panel of experts. Due to the problem of conducting toxicity testing with high doses for hazard assessment <i>in vivo</i> , there is a need to complement animal toxicity studies with a greater understanding of mechanisms and modes of action, to support a dose–response. The experts discussed a) how to improve <i>in vitro</i> systems in order to assess repeated dose toxicity, b) the selection of relevant endpoints for repeated dose toxicity testing, and c) novel biomarkers for

	<p>assessment of toxicity arising from the range of omics technologies. An approach to predict NOEL/NOAEL levels from <i>in vitro</i> methods was proposed and uses the most sensitive endpoints indicative of adverse effects for determination of the NOEC <i>in vitro</i> and biokinetic modelling. A summary of the <i>in vitro</i> models available for five of the most common targets for toxicity (liver, kidney, CNS, lung and haematopoietic system) was also provided. The <i>in vitro</i> systems and endpoints used need to be carefully selected for their <i>in vivo</i> relevance to general mechanisms, as well as to cell-type-specific mechanisms of toxicity or expression of toxic effects.</p>
<p>Grindon, C., Combes, R., Cronin, M.T.D., Roberts, D.W., Garrod, J.F., 2008. An Integrated Decision-tree Testing Strategy for Repeat Dose Toxicity with Respect to the Requirements of the EU REACH Legislation. ATLA 36, 93–101</p>	<p>This paper focuses on the use of alternative (non-animal) methods (both <i>in vitro</i> and <i>in silico</i>) for repeat dose (sub-acute, sub-chronic and chronic) toxicity testing. It reviews the limited number of <i>in silico</i> and <i>in vitro</i> tests available for this endpoint, and outlines new technologies which could be used in the future, e.g. the use of biomarkers and the 'omics' technologies. An integrated testing strategy is proposed, which makes use of as much non-animal data as possible, before any essential <i>in vivo</i> studies are performed. Although none of the non-animal tests are currently undergoing validation, their results could help to reduce the number of animals required for testing for repeat dose toxicity.</p>
<p>Adler, S., Basketter, D., Creton, S., et al., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch. Toxicol. 85(5):367-485.</p>	<p>This report analyses the status and prospects of alternative methods and provides a scientifically sound estimate of the time necessary to achieve full replacement of animal testing in five toxicological areas, i.e. toxicokinetics, repeated dose toxicity, carcinogenicity, skin sensitisation, and reproductive toxicity. For the systemic toxicological endpoint of repeated dose toxicity the time horizon for full replacement could not be estimated. <i>In vitro</i> models in relation to six of the most common targets for repeated dose toxicity (liver, kidney, central nervous system, lung, cardiovascular and haematopoietic system) are included. <i>In silico</i> tools such as (Q)SARs for predicting repeated dose toxicity are also discussed. It was concluded that alternative methods have been developed mainly with the aim of producing stand-alone methods predicting effects in specific target organs. However, for the purpose of quantitative risk assessment better and more scientific knowledge on exposure, toxicokinetics, dose response and mechanisms of toxicity are needed. Additional efforts are necessary to develop improved biokinetic models which would be able to correctly estimate the impact of the distribution over time and level of the repeated external exposure and resulting internal dose. Such models are also needed for extrapolating from <i>in vitro</i> to <i>in vivo</i> and for understanding of dose responses, so that the <i>in vitro</i> data can be applied for quantitative risk assessment. Optimal use of existing data by the</p>

	Threshold of Toxicological Concern concept, read-across and integrated testing strategies can provide an opportunity to avoid the need for <i>in vivo</i> testing for a range of substances and applications.
Basketter, D.A., Clewell, H., Kimber, I., Rossi, A., Blaauboer, B., Burrier, R., Daneshian, M., Eskes, C., Goldberg, A., Hasiwa, N., Hoffmann, S., Jaworska, J., Knudsen, T.B., Landsiedel, R., Leist, M., Locke, P., Maxwell, G., McKim, J., McVey, E.A., Ouédraogo, G., Patlewicz, G., Pelkonen, O., Roggen, E., Rovida, C., Ruhdel, I., Schwarz, M., Schepky, A., Schoeters, G., Skinner, N., Trentz, K., Turner, M., Vanparys, P., Yager, J., Zurlo, J., Hartung, T. 2012. A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report. ALTEX 29(1), 3-91.	This report proposes a roadmap for how to overcome the acknowledged scientific gaps for the full replacement of systemic toxicity testing using animals. Toxicokinetics, skin sensitization, repeated-dose toxicity, carcinogenicity, and reproductive toxicity testing were addressed. For repeated dose toxicity testing the following was recommended: 1) to gather all current data on a wide variety of compounds to improve the quality and speed of new test development and validation; 2) to refine and reduce the number of animals used today by implementing decision trees, tiered approaches and/or applying screening strategies. Existing animal data, data from <i>in vitro</i> tests and data from <i>in silico</i> systems, as well as human data (epidemiological and clinical/medical), can be integrated into these types of approaches; 3) to better understand the molecules and pathways involved in toxicological events (quantification of the alternation of normal signals both the duration or the magnitude of the response); 4) to choose appropriate endpoints for each test and test system; 5) to evaluate the actual toxic dose by considering the real free concentration and the stability of the compounds during the exposure <i>in vitro</i> ; 6) to create a reference list of compounds with information on mechanisms of toxicity and potency readily available in order to speed the validation process for all new testing systems; 7) to be sure of the quality of data used to build <i>in silico</i> models.
Blaauboer, B.J., Boekelheide, K., Clewell, H.J., Daneshian, M., Dingemans, M.M., Goldberg, A.M., Heneweer, M., Jaworska, J., Kramer, N.I., Leist, M., Seibert, H., Testai, E., Vandebriel, R.J., Yager, J.D., Zurlo, J., 2012. The use of biomarkers of toxicity for integrating <i>in vitro</i> hazard estimates into risk assessment for humans. ALTEX 29(4): 411-25.	This report presents the results of a workshop aimed at better defining the use of <i>in vitro</i> -derived biomarkers of toxicity (BoT) and determining the place these data can have in human risk assessment. The selection of BoT takes into account that they need to distinguish adverse and adaptive changes in cells. One of the challenges today when interpreting <i>in vitro</i> data is indeed to distinguish between adversity and adaptation. A framework is presented that defines the place of <i>in vitro</i> systems in the context of data on exposure, structural and physico-chemical properties, and toxicodynamic and biokinetic modelling. The authors highlight the importance of collecting and integrating existing and new data. The selection of these data and their integration needs to be driven by their usefulness in a quantitative <i>in vitro-in vivo</i> extrapolation. It is also important to categorise biomarkers of toxicity and their relation to pathways of toxicity.
Boekelheide K & Schuppe-Koistinen I (2012). SOT/EUROTOX debate:	This article discusses the uses and the limitations of having a gold standard, how

<p>biomarkers from blood and urine will replace traditional histopathological evaluation to determine adverse responses. <i>Toxicological Sciences</i> 129: 249–255.</p>	<p>adverse responses are determined, the evolutionary (as opposed to revolutionary) process by which one technology is typically replaced by another, and the overall goal of developing biomarkers which can translate from preclinical safety assessment to clinical utility. The ultimate purpose of the discussion is to help researchers and regulators understand the challenges they face in the development and integration of new and existing biomarkers to determine adverse responses.</p>
<p>QSARs</p>	
<p>Munro IC, Ford RA, Kennepohl E, Sprenger JG, 1996. Correlation of structural class with No-Observed-Effect Levels: A proposal for establishing a threshold of concern. <i>Food Chem Toxicol</i> 34: 829-867.</p>	<p>This is the key reference providing the basis for relating TTC values to the Cramer Classes. The relationship between chemical structure and toxicity was explored through the compilation of a large reference database consisting of over 600 chemical substances tested for a variety of endpoints resulting in over 2900 no-observed-effect levels (NOELs). Each substance in the database was classified into one of three structural classes using a decision tree approach. The resulting cumulative distributions of NOELs for each of the structural classes differed significantly from one another, supporting the contention that chemical structure defines toxicity. The database was used to derive a threshold of acceptable human exposure for each of the structural classes that could be applied in the absence of specific toxicity data on a substance within one of the three structural classes. The human exposure thresholds provide guidance on the degree of testing and evaluation required for substances that lack toxicity data.</p>
<p>Tilaoui L., Schilter B., Tran L.-A., Mazzatorta P., Grigorov M., 2007. Integrated Computational Methods for Prediction of the Lowest Observable Adverse Effect Level of Food-Borne Molecules. <i>QSAR Comb. Sci.</i>, 26, 102.</p>	<p>The authors present an integrated system partly based on the commercially available software TOPKAT, which predicts chronic toxicity through provision of a computational estimation of Lowest Observed Adverse Effect Level (LOAEL) values. They provide evidence that the LOAEL is correlated with a specific class of molecular descriptors, known as 2D autocorrelation descriptors. The authors concluded that the system developed is helpful in supporting the prioritisation of issues in chemical food research, by establishing levels of safety concern in the absence of sufficient experimental toxicological data.</p>
<p>Lapenna S, Fuart-Gatnik M & Worth A, 2010. Review of QSAR models and software tools for predicting acute and chronic systemic toxicity. JRC Scientific and Technical report EUR 24639 EN. Publications Office of the European Union. Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This JRC report provides a review of different computational estimation methods for predicting acute and chronic systemic toxicity. It provides an overview of Quantitative Structure-Activity Relationship (QSAR) models published in the literature, commonly used software tools, and available databases suitable for QSAR analysis. It also briefly explains the Threshold of Toxicological Concern (TTC) concept and how this is used in prioritising chemicals for further assessment and preliminary risk characterisation.</p>
<p>Spreafico M., Smiesko M., Peristera O., Rossato G., Vedani A., 2010.</p>	<p>The authors describe the addition of a model for LXR to their commercial</p>

<p>Probing Small-Molecule Binding to the Liver-X Receptor: A Mixed-Model QSAR Study, Mol. Info. 29, 27.</p>	<p>software, VirtualToxLab, a fully automated technology that allows for the identification of the endocrine-disrupting potential of drugs, chemicals and natural products. This protocol was applied to screen a series of 161 natural compounds probing their binding to the LXR. The results of the simulation were compared with experimental data (where available), suggesting that the LXR model can be applied to predict the binding affinity of existing or hypothetical compounds for screening purposes. The binding of 52 ligands towards the liver X receptors (LXRs) was identified through docking to the three-dimensional protein structure and quantified by multidimensional QSAR (mQSAR), an approach referred to as 'mixed-model QSAR'. The model was validated by the prediction of 17 external compounds (oxysterols) present neither in the training nor in the test set..</p>
<p>Myshkin E., Brennan R., Khasanova T., Sitnik T., Serebriyskaya T., Litvinova E., Guryanov A., Nikolsky Y., Nikolskaya T., Bureeva S., 2012. Prediction of Organ Toxicity Endpoints by QSAR Modeling Based on Precise Chemical-Histopathology Annotations. Chem. Biol. Drug Des., 80(3), 406.</p>	<p>Under the guidance of the MetaTox consortium (Thomson Reuters, CA, USA), which comprised toxicologists from the pharmaceutical industry and government agencies, the authors created a comprehensive ontology of toxic pathologies for 19 organs, classifying pathology terms by pathology type and functional organ substructure. By manual annotation of full-text research articles, the ontology was populated with chemical compounds causing specific histopathologies. Annotated compound-toxicity associations defined histologically from rat and mouse experiments were used to build QSAR models predicting subcategories of liver and kidney toxicity: liver necrosis, liver relative weight gain, liver lipid accumulation, nephron injury, kidney relative weight gain, and kidney necrosis. All models were validated using two independent test sets and demonstrated overall good performance: initial validation showed 0.80-0.96 sensitivity (correctly predicted toxic compounds) and 0.85-1.00 specificity (correctly predicted non-toxic compounds). Later validation against a test set of compounds newly added to the database in the 2 years following initial model generation showed 75-87% sensitivity and 60-78% specificity. General hepatotoxicity and nephrotoxicity models were less accurate, as expected for more complex endpoints.</p>
TTC	
<p>Lapenna S, Worth A, 2011. Analysis of the Cramer classification scheme for oral systemic toxicity - implications for its implementation in Toxtree. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This JRC report analyses the application of the Threshold of Toxicological Concern (TTC) concept to non-cancer endpoints. The decision tree proposed by Cramer, Ford and Hall in 1978, commonly referred to as the Cramer scheme, is probably the most widely used approach for classifying and ranking chemicals according to their expected level of oral systemic toxicity. The decision tree</p>

	<p>categorises chemicals, mainly on the basis of chemical structure and reactivity, into three classes indicating a high (Class III), medium (Class II) or low (Class I) level of concern. Each Cramer class is associated with a specified human exposure level, below which chemicals are considered to present a negligible risk to human health. In the absence of experimental hazard data, these exposure threshold (TTC) values have formed the basis of priority setting in the risk assessment process. To facilitate the application of the TTC approach, the original Cramer scheme, and an extended version, have been implemented in Toxtree, a freely available software tool for predicting toxicological effects and mechanisms of action. The report provides some suggestions for improving the Cramer scheme based on a review of the scientific literature, a survey of Toxtree users, and an analysis of lists of body and food components incorporated in Toxtree.</p>
<p>EFSA, 2012a. Scientific Opinion on exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC). <i>EFSA Journal</i> 2012;10(7):2750. European Food Safety Authority. Available from: http://www.efsa.europa.eu/en/efsajournal/doc/2750.pdf</p>	<p>EFSA's Scientific Committee developed an opinion on the applicability of the threshold of toxicological concern (TTC) approach as a tool for providing scientific advice about possible human health risks from low level exposures of chemicals in food and feed. The Scientific Committee examined the published literature on the TTC approach, undertook its own analyses and commissioned an <i>in silico</i> investigation of the databases underpinning the TTC approach. The Scientific Committee concluded that the TTC approach can be recommended as a useful screening tool either for priority setting or for deciding whether exposure to a substance is so low that the probability of adverse health effects is low and that no further data are necessary. The following human exposure threshold values are sufficiently conservative to be used in EFSA's work; 0.15 µg/person per day for substances with a structural alert for genotoxicity, 18 µg/person per day for organophosphate and carbamate substances with anti-cholinesterase activity, 90 µg/person per day for Cramer Class III and Cramer Class II substances, and 1800 µg/person per day for Cramer Class I substances, but for application to all groups in the population, these values should be expressed in terms of body weight, i.e. 0.0025, 0.3, 1.5 and 30 µg/kg body weight per day, respectively. Use of the TTC approach for infants under the age of 6 months, with immature metabolic and excretory systems, should be considered on a case-by-case basis. The Committee defined a number of exclusion categories of substances for which the TTC approach would not be used.</p>
<p>EFSA, 2012b. Scientific Opinion on evaluation of the toxicological relevance of pesticide metabolites for dietary risk assessment. <i>EFSA Journal</i> 2012;10(07):2799. European Food Safety Authority. Available</p>	<p>EFSA's Panel on Plant Protection Products and their Residues developed an opinion on approaches to evaluate the toxicological relevance of metabolites and degradates of pesticide active substances in dietary risk assessment. The opinion</p>

<p>from: http://www.efsa.europa.eu/en/efsajournal/doc/2799.pdf</p>	<p>identifies the threshold of toxicological concern (TTC) concept as an appropriate screening tool. The TTC values for genotoxic and toxic compounds were found to be sufficiently conservative for chronic exposure, as a result of a validation study with a group of pesticides belonging to different chemical classes. Three critical steps were identified in the application of a TTC scheme: 1) the estimate of the level of the metabolite, 2) the evaluation of genotoxicity alerts and 3) the detection of neurotoxic metabolites. Tentative TTC values for acute exposure were established by the PPR Panel by analysis of the lowest 5th percentiles of No Observed Adverse Effect Levels (NOAELs) used to establish the Acute Reference Doses (ARfD) for the EFSA pesticide data set. Assessment schemes for chronic and acute dietary risk assessment of pesticide metabolites, using the TTC approach and combined (Q)SAR and read across, are proposed. The opinion also proposes how the risk assessment of pesticide metabolites that are stereoisomers should be addressed due to isomer ratio changes reflected in the composition of metabolites. The approach is ready for use, but it is anticipated that on many occasions the outcome of the assessment scheme will be that further testing is needed to reach a firm conclusion on the toxicological relevance of the metabolite. However, the benefit of applying the approach is that it will allow prioritisation of metabolites for subsequent testing.</p>
<p>SCHER/SCCS/SCENIHR, 2012. Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products. Available from: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_092.pdf</p>	<p>This the European Commission's non-food scientific committees' opinion on the use of the TTC approach for risk assessment of chemical substances in cosmetics and consumer products. They concluded that the TTC approach is interesting as an alternative to costly and time consuming toxicological testing. However, there are two conditions which must be met for TTC to be used as a substitute: 1) where there is little or no information on the toxicity of a chemical substance and 2) where the human exposure is so low that adverse effects are very unlikely to happen. The opinion identifies classes of chemicals and toxic effects for which the TTC approach may be appropriate and those for which it may not. They also addressed the issue of data gaps and research efforts required to strengthen the TTC approach.</p>
<p>Current initiatives to develop alternative approaches</p>	
<p>SEURAT-1 annual report, 2011. Towards the Replacement of in vivo repeated Dose Systemic Toxicity Testing. Volume 1. Available from: http://www.seurat-1.eu/media/download_gallery/SEURAT-1_Annual%20Report_Vol%201_Sept2011_HR.pdf</p>	<p>This publication is the first volume of a series of six Annual Reports that summarises the activities of a new Research Initiative in the field of repeated dose systemic toxicity, which started on 1 January 2011. The aim of the SEURAT-1 Research Initiative is the development of a concept and corresponding long-term research strategy for future research and development</p>

	<p>work leading to pathway based human safety assessments in the field of repeat dose systemic toxicity testing of chemicals. This first Annual Report presents a comprehensive overview of the work in the different projects of the SEURAT-1 Research Initiative. This is given in the context of recent developments in European legislation regarding regulation of chemicals to improve their safety assessment and related international activities. This Research Initiative is designed as a coordinated cluster of five research projects, supported by a ‘data handling and servicing project’ and a ‘coordination and support project’. Detailed project descriptions, reports about the kick-off meetings of each of the specific projects as well as of the SEURAT-1 Research Initiative as a whole are also included. The projects of this research initiative are the following:</p> <ul style="list-style-type: none"> • SCR&Tox: Stem cell differentiation for providing human-based organ specific target cells to assay toxicity pathways <i>in vitro</i>. • HemiBio: Development of a hepatic microfluidic bioreactor mimicking the complex structure and function of the human liver. • DetetcIVE: Identification and investigation of human biomarkers in cellular models for repeated dose <i>in vitro</i> testing. • COSMOS: Delivery of an integrated suite of computational tools to predict the effects of long-term exposure to chemicals in humans based on <i>in silico</i> calculations. • NOTOX: Development of systems biological tools for organotypic human cell cultures suitable for long term toxicity testing and the identification and analysis of pathways of toxicological relevance. • ToxBank: Data management, cell and tissue banking, selection of reference compounds and chemical repository • COACH: Cluster level coordinating and support action
<p>SEURAT-1 annual report, 2012. Towards the Replacement of in vivo repeated Dose Systemic Toxicity Testing. Volume 2. Available from: http://www.seurat1.eu/media/download_gallery/SEURAT%201v2(BD).pdf</p>	<p>This publication is the second volume of a series of six Annual Reports that summarises the activities of the SEURAT-1 Research Initiative in the field of repeated dose systemic toxicity. The core of this second Annual Report is formed by a comprehensive overview of the first results obtained in the different projects of the SEURAT-1 Research Initiative. This is given in the context of recent developments in European legislation regarding the regulation of chemicals to improve their safety assessment and related international activities. The SCR&Tox report focuses on the development of quality control standards that can be applied in routine pluripotent stem cell-based toxicity testing.</p> <ul style="list-style-type: none"> • HeMiBio reports on the generation of the different bioreactor prototypes including the incorporation of high-resolution fluorescent markers into

	<p>pluripotent stem cells as well as the initial complement of electrochemical sensors.</p> <ul style="list-style-type: none"> • DETECTIVE will deliver functional as well as ‘-omics’ biomarkers for different organs (liver, heart and kidney) and reports about first experiments to evaluate the most appropriate human cellular model system for each organ. • Data sets likely to be of use to the COSMOS project, suitable for the development of <i>in silico</i> models, have been identified. A non-cancer dataset about Thresholds for Toxicological Concern (TTC) for cosmetic ingredients has been compiled and the applicability of the TTC approach to cosmetic ingredients has been explored. Furthermore, a process based model able to simulate the dynamics of a chemical compound in cell-based assays has been developed as a basis for <i>in vitro</i> – <i>in vivo</i> extrapolations. • NOTOX describes the establishment of a spheroid cultivation system and its successful application in several cell lines. First toxicity tests have been carried out and ‘-omics’ profiles as well as structural changes were monitored. Tools of integrative and predictive computational systems biology were applied to integrate information obtained from iterative cycles of model predictions and experimental validations by <i>in vitro</i> experiments, in order to eventually predict the possible toxicity of test compounds <i>in vivo</i>. • The ToxBank report focuses on the selection of standard reference compounds (‘gold compounds’) for toxicity testing and the development of a shared cross-cluster database to enable an integrated data analysis. • The COACH report provides information about the cross-cluster coordination, facilitating exchange activities between the projects, and dissemination of research activities at the cluster level.
<p>SEURAT-1, 2013. Towards the Replacement of <i>in vivo</i> repeated Dose Systemic Toxicity Testing. Volume 3. Available from: http://www.seurat-1.eu/media/download_gallery/SEURAT%201-v3%28bd.pdf.</p>	<p>This publication is the third volume of a series of six Annual Reports that summarise the activities of the SEURAT-1 Research Initiative. This third Annual Report, prepared by the coordination and support action project COACH, presents a comprehensive overview of research highlights from the different projects of the SEURAT-1 Research Initiative. This is given in the context of recent developments in European legislation regarding the regulation of chemicals to improve safety assessment and related international activities.</p>
<p>Hepatotoxicity</p>	

<p>Russmann, S., Kullak-Ublick, G.A., and Grattagliano, I. 2009. Current Concepts of Mechanisms in Drug-Induced Hepatotoxicity. <i>Current Medicinal Chemistry</i> 16, 3041-3053</p>	<p>This review summarises current mechanistic concepts of Drug-induced liver injury (DILI) in a 3-step model that limits its principle mechanisms to three main ways of initial injury, i.e. direct cell stress, direct mitochondrial impairment, and specific immune reactions. Initial injury initiates further downstream events, i.e. direct and death receptor-mediated pathways leading to mitochondrial permeability transition, which then results in apoptotic or necrotic cell death. For all mechanisms, mitochondria play a central role in events leading to apoptotic versus necrotic cell death.</p>
<p>Gómez-Lechón, M.J., Lahoz, A., Gombau, L., Castell, J.V., Donato M.T., 2010. In vitro evaluation of potential hepatotoxicity induced by drugs. <i>Curr Pharm Des.</i> 16(17), 1963-77.</p>	<p>This paper describes the molecular mechanisms of hepatotoxicity. All of them have direct effects on organelles such as mitochondria, the endoplasmic reticulum, the cytoskeleton, microtubules, or the nucleus or indirect effects through the activation or inhibition of signalling kinases, enzymes, transcription factors and gene expression profiles. The resulting intracellular stress leads to death by either apoptosis or necrosis. The most important systems to study hepatotoxicity and metabolic activity <i>in vitro</i> are liver slices, isolated liver cells in suspension or in primary cultures including co-culture methods and special 3D techniques, various subcellular fractions and cell lines. These models can be used for cytotoxicity and genotoxicity screening and also to identify the mechanisms involved in drug-induced hepatotoxicity. The development of robust <i>in vitro</i>-based multiparametric screening assays covering a wider spectrum of key effects will improve the predictive capacity for human hepatotoxicity.</p>
<p>Landesmann B., Goumenou M., Munn S., Whelan M. 2012. Description of prototype modes-of-action related to repeated dose toxicity. JRC Scientific and Policy Report EUR 25631 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This report presents the definition and detailed documentation of chosen toxicological MoAs associated with repeated dose target organ toxicity as a first step in building a "prototype" safety assessment framework. In addition to providing a detailed description of the two chosen MoAs related to chronic liver toxicity, namely "MoA from Protein Alkylation to Liver Fibrosis" and "MoA from Liver X Receptor Activation to Liver Steatosis", the report also describes the working process leading to this result including the problems that have been encountered, such as scarcity of quantitative data and the difficulty in capturing and describing complex non-linear processes in a narrative manner. The exercise followed as far as possible relevant WHO-IPCS and OECD guidance and the results have been introduced into a Wiki-based forum that is being developed by the US EPA and the JRC.</p>
<p>Porceddu, M., Buron, N., Roussel, C., Labbe, G., Fromenty, B., and Borgne-Sanchez, A., 2012. Prediction of Liver Injury Induced by Chemicals in Human With a Multiparametric Assay on Isolated Mouse</p>	<p>In this study the authors aimed to improve early prediction of DILI in humans by investigating drug-induced mitochondrial dysfunction as this toxic effect is a major mechanism of DILI. They developed a high-throughput screening platform</p>

<p>Liver Mitochondria. Toxicological Sciences 129(2), 332-345.</p>	<p>using isolated mouse liver mitochondria. Their multiparametric assay was designed to detect the global mitochondrial membrane permeabilization (swelling), inner membrane permeabilization (transmembrane potential), outer membrane permeabilization (cytochrome c release), and alteration of mitochondrial respiration driven by succinate or malate/glutamate. This screening assay revealed high sensitivity for clinical outcome of DILI (94 or 92% depending on cutoff) and a high positive predictive value (89 or 82%). A highly significant relationship between drug-induced mitochondrial toxicity and DILI occurrence in patients was calculated ($p < 0.001$). Moreover, it allowed identifying several compounds for which mitochondrial toxicity had never been described before and even helped to clarify mechanisms with some drugs already known to be mitochondriotoxic. The authors conclude that investigation of drug-induced loss of mitochondrial integrity and function with this multiparametric assay should be considered for integration into basic screening processes at an early stage to select drug candidates with lower risk of DILI in human. This assay is also a tool for assessing the mitochondrial toxicity profile and investigating the mechanism of action of new compounds and marketed compounds.</p>
<p>Tolosa, L., Pinto, S., Donato, M.T., Lahoz, A., Castell, J.V., O'Connor, J.E., and Gómez-Lechón, M.J., 2012. Development of a Multiparametric Cell-based Protocol to Screen and Classify the Hepatotoxicity Potential of Drugs. Toxicological Sciences 127(1), 187-198.</p>	<p>In this paper the authors aimed to develop a practical, reproducible, <i>in vitro</i> multi-parametric cell-based protocol to assess those drugs that are potentially hepatotoxic to humans and to suggest their mechanisms of action. They used HepG2 human cell line cultured in 96-well plates and exposed to 78 different compounds for 3 and 24 h at different concentrations. They measured parameters associated with nuclear morphology, plasma membrane integrity, mitochondrial function, intracellular calcium concentration, and oxidative stress, indicative of pre-lethal cytotoxic effects and representative of different mechanisms of toxicity. These parameters were measured at the single cell level with the high-content screening (HCS) technology, which allows high-throughput screening. The strategy presented appears to identify early and late events in the hepatotoxic process and also suggests the mechanism(s) implicated in the toxicity of compounds to thereby classify them according to their degree of injury (no injury, low, moderate, and high injury).</p>
<p>Vinken M, Landesmann B, Goumenou M, Vinken S, Shah I, Jaeschke H, Willett C, Whelan M, Rogiers V, (2013). Development of an adverse outcome pathway from drug-mediated bile salt export pump inhibition to cholestatic liver injury. Toxicol Sci. 136(1): 97-106.</p>	<p>Cholestasis insults are among the most severe clinical manifestations of drug-induced liver injury. In this article, an Adverse outcome pathway (AOP) framework is proposed for cholestasis triggered by drug-mediated inhibition of the bile salt export pump transporter protein. An in-depth survey of relevant scientific literature was carried out in order to identify intermediate steps and key events. The latter include bile accumulation, the induction of oxidative stress and</p>

	<p>inflammation, and the activation of specific nuclear receptors. Collectively, these mechanisms drive both a deteriorative cellular response, which underlies directly caused cholestatic injury, and an adaptive cellular response, which is aimed at counteracting cholestatic insults. The postulated AOP, which was developed according to the Organisation for Economic Co-operation and Development (OECD) guidance, is expected to serve as the basis for the development of new <i>in vitro</i> tests and the characterization of novel biomarkers of drug-induced cholestasis.</p>
<p>Cellular pathways</p>	
<p>Wilmes, A., Crean, D., Aydin, S., Pfaller, W., Jennings, P., Leonard, M.O., 2011. Identification and dissection of the Nrf2 mediated oxidative stress pathway in human renal proximal tubule toxicity. <i>Toxicology in Vitro</i> 25, 613-622.</p>	<p>The authors have used whole genome DNA microarrays in an attempt to uncover molecular mechanisms of response to nephrotoxin exposure. For this purpose they treated human proximal tubular epithelial cells (HK-2 cell line) with three unrelated chronic nephrotoxins: the heavy metal and environmental pollutant cadmium, the contact herbicide diquat dibromide and the immunosuppressive calcineurin inhibitor cyclosporine. Using bioinformatic pathway analysis they identified the Nrf2 pathway as the most prominently altered pathway. Furthermore, they could demonstrate a Nrf2 dependent induction of HO-1 and NQO1 and that HO-1 is protective to cadmium toxicity. This study highlights the importance of HO-1 as a possible common biomarker for nephrotoxicity and as a mechanism of protection against nephrotoxic chemical damage.</p>
<p>Jennings, P., Limonciel, A., Felice, L., Leonard, M.O. 2013. An overview of transcriptional regulation in response to toxicological insult. <i>Arch. Toxicol.</i> 87:49-72.</p>	<p>This review summarises the major toxicologically relevant transcription factor-governed molecular pathways. The authors focus on current knowledge of transcriptional responses to toxic exposure and describe how toxins can directly or indirectly produce distinct gene expression bio-signatures. They focus on the top-level response of transcription factor activation:</p> <ul style="list-style-type: none"> • The cap ‘n’ collar (CNC) subfamily of bZIP transcription factors, which is one of the best characterised responders to toxin exposure and is activated in response to oxidative or electrophilic stress. • The p53 pathway, which plays a pivotal role in differentiation and tumour suppression and has been extensively studied in many contexts including irradiation, carcinogenicity, ageing and toxicity. • The aryl hydrocarbon receptor which is a member of the bHLH/PAS family of transcription factors highly conserved through evolution and initially characterised as mediating the toxic effects of pollutants. • The nuclear receptor family which is the largest group of transcriptional regulators involved in xenobiotic sensing and toxicological response.

	<ul style="list-style-type: none"> • Immunomodulatory transcription factors: interference with normal immune regulatory signalling pathways is a primary cause of immunotoxicity observed with xenobiotic exposure. • Hypoxia-inducible factors: its activation orchestrates the transcriptional regulation of many processes including, erythropoiesis, glycolysis and angiogenesis. • The unfolded protein response which is mounted when ER stress is detected and serves primarily to return normal endoplasmic reticulum function. • Metal responsive transcription factor-1, which is primarily involved in directing gene expression in response to various heavy metals but can also respond to other cellular stresses such as hypoxia and oxidative chemical species. • Heat shock factors (HSF1-4) are considered major responders to proteotoxic stimuli including heat shock for which they were first discovered, and diverse environmental chemical and physiological stresses where protein mis-folding is the common factor. • Mitochondria is a major target for toxicant induced cellular injury, and alterations in the transcriptional regulation of components of this important organelle are clear indications of cellular stress. • Tissue-specific transcriptional regulation and in particular the authors discuss transcription factor control of renal epithelial cell differentiation and the response to toxicant-induced injury. <p>The authors conclude that it is possible to gain important information regarding possible mechanisms of toxicity from analysing the likely transcriptional regulators responsible for the pattern of gene expression. This can now be achieved through the use of pathway analysis software.</p>
<p>Wilmes, A., Limonciel, A., Aschauer, L., Moenks, K., Bielow, C., Leonard, M.O., Hamon, J., Carpi, D., Ruzek, S., Handler, A., Schmal, O., Herrgen, K., Bellwon, P., Burek, C., Truisi, G.L., Hewitt, P., Di Consiglio, E., Testai, E., Blaauboer, B.J., Guillou, C., Huber, C.G., Lukas, A., Walter, P., Mueller, S.O., Bois, F.Y., Dekant, W., Jennings, P., 2013. Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. <i>J Proteomics</i> 79, 180-194.</p>	<p>This paper investigates the added benefit of integrating transcriptomics, proteomics and metabolomics together with pharmacokinetics for application in drug safety, utilising cultured human renal epithelial cells treated with therapeutic and suprathreshold concentrations of the nephrotoxin Cyclosporine A in repeat dose regimes (14 days). The results clearly showed that cells are simultaneously experiencing mitochondrial perturbations, oxidative- and ER-stress only at the high CsA concentration. However, CyP-B secretion was maximum already at the low CsA concentration. This study demonstrates that <i>in vitro</i> cell culture systems coupled with</p>

	<p>pharmacokinetics and high content omic approaches can give extremely detailed and quantitative insights into both the pharmacological and toxicological effects of compounds.</p>
<p>Hamon J, Jennings P, Bois FY, (2014). Systems biology modeling of omics data: effect of cyclosporine a on the Nrf2 pathway in human renal cells. BMC Syst Biol. 8(1): 76. doi:10.1186/1752-0509-8-76.</p>	<p>One of the challenges of quantitative modeling of toxicity pathways is to integrate omics data with systems biology models for parametric inference and model checking. In this article the authors present a quantitative calibration of a differential equation model of the nuclear factor (erythroid-derived 2)-like 2 pathway (Nrf2) pathway with a subset of the omics data previously published by Wilmes et al. (2013). In that publication Wilmes and colleagues showed that cyclosporine A (CsA) strongly activates Nrf2 in renal proximal tubular epithelial cells (RPTECs) exposed <i>in vitro</i>.</p> <p>Modeling was done into two steps: (i) Modeling the <i>in vitro</i> pharmacokinetics (PK) of CsA (exchange between cells, medium and vial walls) with a minimal distribution model. (ii) Modeling the effects of CsA on omics markers at the cellular level with a coupled PKsystems biology model. Posterior statistical distributions of the model parameter values were obtained by Markov chain Monte Carlo sampling in a Bayesian framework. The authors concluded that the model proposed can be used to analyze and predict cellular response to oxidative stress, provided sufficient data to set its parameters to cell-specific values. Omics data can be used to that effect in a Bayesian statistical framework which retains prior information about the likely parameter values.</p>
<p>Biokinetic considerations</p>	
<p>Zaldivar Comenges J.M., Menecozzi, M., Macko, P., Rodriguez, R., Bouhifd, M., and Baraibar Fentanes, J., 2011. A Biology-Based Dynamic Approach for the Modelling of Toxicity in Cell Assays: Part II: Models for Cell Population Growth and Toxicity. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/</p>	<p>This paper provides a mathematical approach for extrapolating from <i>in vitro</i> concentrations to <i>in vivo</i> dose. To do this extrapolation it is necessary to be able to calculate and compare the free concentrations in both systems</p> <p>Concerning the <i>in vitro</i> side, the authors has previously developed and implemented, based on HTS (High Throughput Screening) laboratory data, a compound fate model using the partitioning approach. The developed fate model was able to predict the role of serum in toxicity assays as well as provide estimation on the partitioning of a certain compound between the headspace, plastic wall and the medium (attached to serum, free dissolved and attached to the cells). However, the partitioning approach assumes that the equilibrium is fast in comparison with the duration of the experiments, which could not be the case for the partitioning to the cells. For this reason, a DEB (Dynamic Energy Budget) stage-based toxicity model was developed and experimentally verified in this work. The results show that this approach offers the possibility of extrapolating</p>

	the values obtained to calculate <i>in vivo</i> human toxicology thresholds using a PBTK modelling approach.
Coecke, S., Pelkonen, O., Leite, S.B., Bernauer, U., Bessems, J.G., Bois, F.Y., Gundert-Remy, U., Loizou, G., Testai, E., Zaldívar, J.M., 2013. Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. <i>Toxicol. In Vitro</i> 27(5): 1570-7..	<p>This review summarises the current situation regarding toxicokinetics and it develops further some critical areas, such as crucial requirements of <i>in vitro</i> kinetics and conditions for the total replacement of animal toxicity tests. A set of recommendations are also provided. The conclusions from the authors are as follows:</p> <ul style="list-style-type: none"> - For a proper <i>in vitro</i>–<i>in vivo</i> extrapolation of toxic responses, it is imperative to understand <i>in vitro</i> kinetics in <i>in vitro</i> toxicity testing systems. - Without the toxicokinetic information it remains questionable whether toxicodynamic effects observed at specific <i>in vitro</i> concentration ranges have any relevance for human exposure condition. - These contributions are a key for a further development of animal-free testing for systemic toxicity. - Full integration of kinetic expertise into design and execution of toxicity testing and risk assessment is essential, however, when risk assessment uses only <i>in vitro</i> data, it is indispensable to link toxicokinetic with toxicodynamic effect models based on <i>in vitro</i> tests. - For several kinetic parameters needed to parameterise PBTK models, non-animal methods are available for chemicals at various levels of development. However, the authors concluded that alternative methods are lacking completely for predicting renal and biliary excretion, as well as absorption in the lung.
Diaz Ochoa JG, Bucher J, Péry AR, Zaldivar Comenges JM, Niklas J, Mauch K. 2012. A multi-scale modeling framework for individualized, spatiotemporal prediction of drug effects and toxicological risk. <i>Frontiers in Pharmacology</i> 3, 204.	<p>This paper describes a multi-scale modelling approach for spatiotemporal prediction of the distribution of substances and resulting hepatotoxicity by combining cellular models, a 2D liver model, and whole body model. As a case study, the authors focused on predicting human hepatotoxicity upon treatment with acetaminophen based on <i>in vitro</i> toxicity data and potential inter-individual variability in gene expression and enzyme activities. By aggregating mechanistic, genome-based <i>in silico</i> cells to a novel 2D liver model and eventually to a whole body model, they predicted pharmacokinetic properties, metabolism, and the onset of hepatotoxicity in an <i>in silico</i> patient. Depending on the concentration of acetaminophen in the liver and the accumulation of toxic metabolites, cell integrity in the liver as a function of space and time as well as changes in the elimination rate of substances were estimated. They showed that the variations in elimination rates also influence the distribution of acetaminophen and its metabolites in the whole body. The integrated model also predicted variations in drug toxicity depending on alterations of metabolic enzyme activities. Variations</p>

	in enzyme activity, in turn, reflect genetic characteristics or diseases of individuals.
Broeders, J.J.W., Blaauboer, B.J., Hermens, J.L.M., 2013. <i>In vitro</i> biokinetics of chlorpromazine and the influence of different dose metrics on effect concentrations for cytotoxicity in Balb/c 3T3, Caco-2 and HepaRG cell cultures. <i>Toxicology in Vitro</i> 27, 1057–1064.	In this paper, the cytotoxicity (measured by Alamar Blue) and biokinetics of the antipsychotic chlorpromazine were studied in <i>in vitro</i> assays using different cell types (Balb/c 3T3, Caco-2 and HepaRG cells) and exposure conditions. Different dose metrics (the nominal, measured total and measured free chlorpromazine medium concentrations) were assessed to express the sensitivity to chlorpromazine. Based on the results the authors had the following observations: (i) the concentration in the medium may decrease over time simply due to uptake into the cells, (ii) the decrease in medium concentration will be higher in an <i>in vitro</i> test with a higher cell density, (iii) the decrease in medium concentration over time will be lower at dose levels where cytotoxicity occurs, and (iv) conclusions about differences in sensitivity between cell systems based on nominal or even measured total concentrations often do not represent the differences in intrinsic sensitivity, and therefore they concluded that the freely dissolved concentration is a more appropriate dose metric than total concentration in the medium in comparing the sensitivity of the tested cell systems. The ranking in sensitivity was dependent on the dose metric used.
Schug, M., Stöber, R., Heise, T., Mielke, H., Gundert-Remy, U., Godoy, P., Reif, R., Blaszkewicz, M., Ellinger-Ziegelbauer, H., Ahr, H.J., Selinski, S., Günther, G., Marchan, R., Blaszkewicz, M., Sachinidis, A., Nüssler, A., Oberemm, A., Hengstler, J.G., 2013. Pharmacokinetics explain <i>in vivo/in vitro</i> discrepancies of carcinogen-induced gene expression alterations in rat liver and cultivated hepatocytes. <i>Arch. Toxicol.</i> 87, 337-345.	In this study the authors aimed to elucidate the reason why the non-genotoxic rat liver carcinogen methapyrilene alters the expression of the metabolizing genes <i>SULT1A1</i> and <i>ABAT</i> , as well as the DNA damage response gene <i>GADD34</i> <i>in vitro</i> , but not <i>in vivo</i> . There are two major limitations when using cultivated primary hepatocytes in toxicogenomics. Firstly, hepatocytes undergo massive gene expression alterations, particularly during the first 24h in culture. Therefore, gene expression alterations induced by test compounds have to be analyzed against a “noisy” background. Secondly, huge discrepancies between test compound-induced gene expression alterations in the liver <i>in vivo</i> and in hepatocytes <i>in vitro</i> have been reported. If the <i>in vitro</i> rat primary hepatocytes really induce false-positive results, further investment into its development as a tool for toxicogenomics for the <i>in vivo</i> situation is not justified. However, the data of this study demonstrate that the putative discrepancy has a pharmacokinetic explanation. The relatively short half-life of 2.8 h implies a rapid decrease in orally administered methapyrilene <i>in vivo</i> below concentrations that can cause gene expression alterations. RNA levels are altered 1, 6 and 12 h after methapyrilene administration, but return to control levels after 24 and 72 h. In contrast, methapyrilene concentrations in the culture medium supernatant of primary rat hepatocyte cultures decreased slowly. This explains why <i>GADD34</i> ,

<p>ABAT and SULT1A1 were still deregulated after 24 h exposure <i>in vitro</i>, but not <i>in vivo</i>. The different pharmacokinetics in the 6-well dish compared to the <i>in vivo</i> situation is an interesting example to demonstrate the need for so-called <i>in vitro</i> biokinetic studies, in particular, if effects are to be monitored over a time period of several days. The authors conclude that the different pharmacokinetics of methapyrilene HCl <i>in vivo</i> and <i>in vitro</i> sufficiently explains the previously reported discrepancy.</p>
--

8. Genotoxicity and mutagenicity

Raffaella Corvi, Federica Madia & Andrew Worth

8.1 *The traditional in vitro and in vivo tests and their regulatory use*

Genetic alterations in somatic and germ cells are associated with serious health effects, which in principle may occur even at low exposure levels. Mutations in somatic cells may cause cancer if mutations occur in proto-oncogenes, tumour suppressor genes and/or DNA damage response genes, and are responsible for a variety of genetic diseases (Erickson, 2010). Accumulation of DNA damage in somatic cells has also been proposed to play a role in degenerative conditions such as accelerated aging, immune dysfunction, cardiovascular and neurodegenerative diseases (Hoeijmakers, 2009; Slatter & Gennery, 2010; De Flora & Izzotti, 2007; Frank, 2010). Mutations in germ cells can lead to spontaneous abortions, infertility or heritable damage to the offspring and possibly to the subsequent generations.

Currently, the assessment of genotoxic hazard to humans follows a stepwise approach, beginning with a basic battery of *in vitro* tests followed in some cases by *in vivo* testing. Regulatory requirements, in particular for *in vivo* testing, vary depending on the type of chemical under regulation and the region. For cosmetics, *in vivo* testing is prohibited in the EU (EC No 1223/2009; SCCS, 2012), while for industrial chemicals and biocidal products a positive outcome in one or more of the *in vitro* genotoxicity tests requires confirmation by appropriate follow-up *in vivo* testing (EC no. 1907/2006; EC no. 528/2012)). In these cases, if a substance is clearly negative in the *in vitro* battery it is considered as having no genotoxic hazard, thus no further *in vivo* study is needed. Regulatory requirements for pharmaceuticals, veterinary drugs and plant protection products foresee that the *in vitro* testing battery (irrespective of the outcome) is always followed by *in vivo* testing (ICH, 2011; VICH, 2012; EC No 1107/2009 and 283/284-2013).

The standard *in vitro* test battery comprises the bacterial reverse mutation assay (OECD TG 471), the *in vitro* mammalian chromosomal aberration test (OECD TG 473), and the *in vitro* mammalian cell gene mutation test (OECD TG 476), and the *in vitro* mammalian cell micronucleus test (OECD TG 487).

In vivo tests comprise the mammalian erythrocyte micronucleus test (OECD TG 474), the mammalian bone marrow chromosomal aberration test (OECD TG 475), the rodent dominant lethal assay (OECD TG 478), the mammalian spermatogonial chromosome aberration test (OECD TG 483), the mouse heritable translocation assay (OECD TG 485), the unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo* (OECD TG 486), the transgenic rodent somatic and germ cell gene mutation assay (OECD TG 488) and the *in vivo* mammalian alkaline comet assay (draft OECD TG).

The standard genotoxicity tests are summarised in Table 8.1.

8.2 *Mechanistic understanding of the endpoint*

Genotoxicity testing include both the measurement of direct, irreversible damage to the DNA ("mutagenicity"), that is transmissible to the next cell generation, as well as the measurements of early, potentially reversible effects on the DNA or on mechanisms involved in the preservation of the integrity of the genome ("genotoxicity"). The definitions of the terms "mutagenicity" and

"genotoxicity" are taken from the REACH "Guidance on information requirements and chemical safety assessment" (ECHA, 2012).

Mutagenicity refers to the induction of permanent transmissible changes in the amount or structure of the genetic material of cells or organisms. These changes may involve a single or block of base pairs, single gene or gene segment, a block of genes or chromosomes. The term clastogenicity is used for agents giving rise to structural chromosome aberrations. A clastogen can cause breaks in chromosomes that result in the loss or rearrangements of chromosome segments. Aneugenicity refers to the effects of agents that give rise to a change (gain or loss) in chromosome number.

Genotoxicity is a broader term and refers to processes, which alter the structure, information content or segregation of DNA and are not necessarily associated with mutagenicity. Thus, tests for genotoxicity include tests which provide an indication of induced damage to DNA (but not direct evidence of mutation), as well as tests for mutagenicity.

The detection of changes or damage directly to the DNA without measuring the adverse consequences of the genetic damage is generally performed with indicator tests. For example, some covalent DNA adducts do not change the coding properties of the DNA. Other damage such as that measured by the comet assay may be accurately repaired or inevitably leads to cell death, which does not pose a genotoxic hazard to the organism. Indicator tests may provide a measure of (target tissue) exposure or information on the mode of action (e.g. by indicating DNA interaction).

For an adequate evaluation of the genotoxic potential of a chemical substance, different endpoints (i.e. induction of gene mutation, structural and numerical chromosome alterations) have to be assessed, as each of these events has been implicated in carcinogenesis and heritable diseases but an adequate coverage of all endpoints can only be obtained by the use of multiple test systems (i.e. a testing battery), as no individual test system currently covers all endpoints.

Even though a number of well-established *in vitro* methods are available and officially accepted for genotoxic hazard assessment (Table 8.1), new methods and strategies continue to be developed. This is because the existing *in vitro* methods, while having a high sensitivity (and thus low false negative rate), have a relatively low specificity and thus high rate of false ("misleading") positive results, which typically leads to follow-up testing *in vivo* for the confirmation of these results (Kirkland et al., 2005). During a workshop organised by ECVAM (2006, Ispra, Italy) the high rate of unexpected positive results in *in vitro* mammalian cell genotoxicity tests was addressed (Kirkland et al., 2007). It was recommended that better guidance on the likely mechanisms resulting in positive results not relevant for humans, and how to obtain evidence for these mechanisms was needed. Among the identified reasons for misleading positives are high cytotoxicity, high passage number, deficient p53 status, deficient DNA repair. The workshop recommendations paved the way for several international collaborative initiatives aiming to improve the existing genotoxicity *in vitro* tests and to identify and evaluate new cell systems with appropriate sensitivity but improved specificity.

Several options are currently being explored to improve the overall assessment of genotoxicity: improvement of the existing standard tests, development of new tests, and development of integrated testing strategies (ITS) with emphasis on reducing the need for animal testing.

8.3 Improvement of standard tests

The first option is partly addressed by the current revision of the OECD TGs for genotoxicity and development of an Introduction document to the OECD guidelines on genetic toxicology testing and a Guidance document on the selection and application of assays, in support of the Genotoxicity Test Guidelines (in preparation). The revision is taking into account the knowledge acquired during the last decades of testing and the results from recent collaborative studies (Parry et al., 2010; Kirkland and Fowler, 2010; Fowler et al., 2012a, 2012b), including recommendations on testing at reduced maximum concentration, careful choice of the cell lines and measure of cytotoxicity. The revised TGs are expected to enhance the quality of the data produced and consequently avoid in some cases the need for *in vivo* confirmation of the results.

8.4 Development of new *in vitro* tests

A number of new test methods are being developed and validated with the aim of improving the specificity of *in vitro* tests while maintaining an appropriate sensitivity. For the development of new test systems, an ECVAM workshop concluded that cell systems of human origin which are p53 and DNA repair proficient, and have defined phase 1 and phase 2 metabolism, offer the best option to reduce unexpected positive or negative results in the future (Kirkland et al, 2007).

The micronucleus test and the comet assay in 3D human reconstructed skin models offer the potential for a more physiologically relevant approach for dermal exposure and metabolism (Hu et al., 2010). It is anticipated that these reconstructed skin models could improve the predictive value of a genotoxicity assessment compared with existing *in vitro* tests and, therefore, could be used as a follow-up test in case of positive results from the standard *in vitro* genotoxicity testing battery (Maurici et al., 2005; Pfuhler et al., 2010; Kirsch-Volders et al., 2011). Several 3D skin models are commercially available and are suitable for conducting such tests, provided that sufficient cell proliferation is available. A micronucleus test protocol using the EpiDerm™ model (MatTek Corporation, Ashland, MA, USA) has been developed and evaluated with a variety of chemicals across three laboratories (Curren et al., 2006; Mun et al., 2009; Hu et al. 2010). An on-going multi-laboratory prevalidation study coordinated by Cosmetics Europe aims at establishing the reliability of the method (Aardema et al., 2010) and at increasing the domain of chemicals tested for predictive capacity. The comet assay in reconstructed skin models is a rapid and sensitive method to evaluate primary DNA damage and it could be used as a follow-up test for chemicals that cause gene mutation in the *in vitro* standard tests. Indeed, the International Workshop on Genotoxicity Testing (2009) suggested the *in vitro* comet assay in the 3D skin model as a valuable support to genotoxicity identification as i) it is, differently than the micronucleus test, independent of cell proliferation and ii) it covers a broader spectrum of DNA damage. Similarly to the micronucleus test, a protocol using the EpiDerm™ model has been developed and evaluated across different laboratories (Reus et al., 2013). An on-going collaborative study between Cosmetics Europe and a German Consortium funded by BmBF regards also the evaluation of the comet assay in epidermis full thickness: a reconstructed skin model with a fully developed *in vivo*-like basement membrane (EpiDerm-FT™). Ultimately, it will be important to demonstrate whether these tests have an equivalent sensitivity and a better specificity compared to the standard *in vitro* genotoxicity mammalian cell tests.

Another promising system that has been proposed as a follow-up for *in vitro* positives is the hen's egg test for micronucleus induction (HET-MN; Wolf et al., 2008). The HET-MN combines the use of the commonly accepted genetic endpoint “formation of micronuclei” with the well-

characterized model of the incubated hen's egg, which enables metabolic activation, elimination and excretion of xenobiotics, including those that are mutagens or pro-mutagens. Studies on metabolism indicate that certain important phase I and II enzymes are active and, therefore, the detection of liver mutagens is possible. A German Consortium ring trial prevalidation study is on-going and a first report has been published on the reproducibility and robustness of the HET-MN for the prediction of genotoxicity (Greywe et al., 2012).

The *in vitro* cell transformation assays (CTA) are not genotoxicity assays, but they have been established in order to predict carcinogenicity (DiPaolo et al., 1969; Isfort et al., 1996; Matthews et al., 1993; LeBoeuf et al., 1999; SLA on carcinogenicity). The CTAs measure the transformation of cells and have the potential to detect both genotoxic and non-genotoxic carcinogens. The assays could be used as a follow-up assay for confirmation of *in vitro* positive results from genotoxicity assays as part of a weight of evidence assessment. Furthermore, data generated by the CTA can be useful for the genotoxicity assessment of chemical classes that are not reliably predicted by traditional *in vitro* genotoxicity tests (e.g. aromatic amines) (EFSA, 2011; Vanparys et al., 2012; EURL ECVAM, 2012). An OECD TG on CTA in Syrian hamster embryonic stem cells is being drafted (draft OECD TG on SHE CTA). Further information on CTA for the assessment of transforming potential is given in the carcinogenicity chapter.

Toxicogenomics can be used to identify global gene expression changes associated with a toxicological outcome, including carcinogenicity and genotoxicity. In the context of genotoxicity testing, its primary use is envisaged to be in providing information on mode of action and such information can be useful supporting evidence (Doktorova et al., 2012). The application of toxicogenomics to predict mode of action has been recently reviewed in depth (Ellinger-Ziegelbauer et al., 2009; Waters et al., 2010). Although the published *in vitro* and *in vivo* data set show appreciable variability, common features emerge with respect to molecular pathways. For instance, the DNA damage-responsive p53 pathway is extensively activated both by DNA reactive genotoxins *in vitro* and genotoxic carcinogens *in vivo* (Doktorova et al., 2013; Jennings et al., 2013). Conversely, *in vitro* DNA non-reactive genotoxins and *in vivo* non-genotoxic carcinogens mostly induce genes associated with oxidative stress, signalling and cell cycle progression. These data represent a first proof of concept that the gene expression profiles reflect the underlying mechanism of action. However, additional studies should be performed to enlarge the number of chemicals tested, to fill the gaps in dose-response and time-course relationships and in the case of *in vivo* toxicogenomics to analyse different routes of exposure and organ systems (most studies so far have used rat liver) and other species. At the moment, these tests should be used to generate supportive mechanistic information to improve the genotoxicity assessment, rather than for routine testing.

In the last few years several attempts have been made to develop and validate the induction of stress pathways/proteins as endpoints in genotoxicity assays by using high throughput screening approaches. The choice of the pathways was mostly based on microarray experiments with genotoxic substances. The GreenScreen HC assay, which uses a p53-competent TK6 lymphoblastoid cell line genetically modified to incorporate a fusion cassette containing the genotoxic stress-inducible GADD45 α gene sequence including promoter region with the GFP gene as reporter (GADD45 α -EGFP reporter plasmid pEPGD532) (Hastwell et al., 2006), has been widely characterised and its sensitivity and specificity confirmed in independent studies (Hastwell et al., 2009; Jagger et al., 2009; Birrel et al., 2010). This assay was further adapted to test chemicals that require metabolic activation (Billinton et al., 2010; Hughes et al., 2012). HTS luciferase reporter assays based on four different stress pathways (RAD51, Cystatin A p53 and

Nrf2) in the HepG2 cell line have also been developed and shown to be useful for screening in early phases of drug development with the potential to reduce the attrition rate due to genotoxicity in the developmental phase of drug development (Westernik et al., 2010). On the basis of currently available information, these assays represent suitable *in vitro* screening methods for the ability to detect DNA damage induced by a broad spectrum of genotoxic insults.

8.5 *In silico* models

A number of *in silico* tools (QSAR models and expert systems) for genotoxicity are available, including freely available software in the public domain. These tools have been extensively characterised in the scientific literature (Serafimova et al 2010; Worth et al, 2010). In general, available QSAR models are based on DNA reactivity and predict Ames mutagenicity with high sensitivity. As with *in vitro* tests, this is generally accompanied with a high percentage of false positive predictions. Public domain tools include CAESAR (<http://www.vega-qsar.eu/>) and Lazar (Helma, 2006).

Expert systems are ruled-based, with the rules being derived by expert knowledge and/or statistical induction. These systems typically aggregate data from multiple sources, and may therefore provide a “high level” assessment of genotoxic potential rather than reproduce specific and standardised test methods. Public domain tools include rulebases in Toxtree (Benigni et al., 2008) and profilers in the OECD QSAR Toolbox (OECD, 2010d). The QSAR Toolbox can be used to group chemicals for read-across, an approach which may also be used to identify non-genotoxic chemicals (Worth et al., 2012). Relatively few *in silico* tools are available for the prediction of *in vivo* genotoxic potential, and further efforts are needed to develop and evaluate these methods.

8.6 Development of integrated testing strategies

It is important to consider all available, relevant knowledge on the substance which may have an impact upon the selection of tests and/or may indicate that more or fewer genotoxicity tests are needed on the substance than defined by the standard information requirements. This existing information, which may include other toxicity data, physico-chemical and biokinetics properties, confounding factors, as well as mechanistic data, may be helpful for the interpretation of *in vitro* genotoxicity test results (Dearfield et al., 2011). Additional *in vitro* tests may also be considered as part of follow-up testing for *in vitro* positive genotoxic results to generate supplementary information regarding mechanism (or mode) of action and to clarify the relevance of the positive results for humans. In this regard, *in vitro* tests which detect primary DNA damage (defined as indicator tests) or tests that measure gene expression may be considered particularly useful. At the moment the results of these non-standard methods are to be considered to add weight of evidence for improved genotoxicity risk assessments in humans (Lynch et al., 2011).

Recently, the EFSA Scientific Committee and the UK Committee on Mutagenicity of Chemicals in Food, Consumers Products, and the Environment (COM) recommended that the set of tests used in the *in vitro* battery be as small as scientifically justifiable in order to reduce the number of misleading positives (EFSA, 2011; COM, 2011). Based on a recent data analysis (Kirkland et al., 2011), a two tests initial step of testing has been proposed which includes the bacterial reverse mutation assay (OECD TG 471) and the *in vitro* micronucleus assay (OECD TG 487). As the *in vitro* micronucleus assay detects two of the endpoints (i.e. structural and numerical chromosome aberration), the proposed two-test battery covers all three endpoints to be assessed in the standard testing battery.

With the objective of reducing the use of experimental animals, the choice of the *in vivo* follow-up test is critical and needs to cover the same endpoint as the one which showed positive results *in vitro* (Kirkland & Speit, 2008). As a default only one *in vivo* test should be performed initially. A second *in vivo* test should only be considered if the first *in vivo* test is negative, and does not cover all *in vitro* positive genotoxic endpoints (Dearfield et al., 2011). New *in vivo* tests have recently been developed in order 1) to be able to evaluate genotoxicity in almost all tissues, and 2) to measure endpoints other than chromosome damage. However, with an increased number of *in vivo* tests available, there is a risk that more animals could be used in the future for the assessment of genotoxicity. Therefore, selection of the appropriate follow-up test is particularly important in view of the adoption of new *in vivo* genotoxicity OECD TGs (OECD TG 488 and draft OECD TG on comet assay).

Several opportunities for reduction exist both at single test level (e.g. 1 sex versus 2 sexes, smaller animal groups) and ITS level (Pfuhrer et al., 2009; EFSA, 2011). The integration of different endpoints into a single study (Pfuhrer et al., 2009; Bowen et al., 2011) or the incorporation of *in vivo* genotoxicity endpoints into a short-term repeated dose toxicity test (28 days) (Rothfuss et al., 2010, 2011; EFSA, 2011), if such a test is going to be performed anyhow, should always be considered. Most of the currently accepted *in vivo* tests are amenable to such integration. An integration of genotoxicity endpoints offers the possibility for an improved interpretation of genotoxicity findings since such data will be evaluated in conjunction with routine toxicological information obtained in the repeated dose toxicity study, such as haematology, clinical chemistry, histopathology and exposure data (Pfuhrer et al., 2009). For substances that are somatic genotoxicants *in vivo*, the potential to affect germ cells should also be investigated. It is reasonable to conclude that the substance can cause heritable genetic damage, provided that it has the ability to reach the gonads as its probably to interact with the DNA would be very high. Then no further testing should be performed and the substance can be considered a germ cell mutagen.

8.7 Conclusions

Several *in vitro* tests are available and officially accepted for the assessment of genotoxicity. However, despite the high sensitivity of these methods, the high percentage of false positives (low specificity) is problematic, and in most regulatory frameworks (including REACH) this triggers follow-up *in vivo* testing (Adler et al., 2011).

A number of *in silico* (QSAR) tools are also available for predicting genotoxicity, and in particular *in vitro* mutagenicity. These tools can be used in a weight of evidence approach to add confidence to *in vitro* test results.

A range of initiatives are aiming at the development of a new generation of *in vitro* tests, but none is yet sufficiently well developed and standardised for routine application in regulatory testing. In certain cases though, the results from these tests may be useful to provide supplementary data in a weight of evidence approach.

The most promising near-term opportunities for applying the 3Rs in genotoxicity assessment will be through the improvement of standard tests and the development of integrated testing strategies with emphasis on reduction alternatives (EURL ECVAM, 2013).

Table 8.1. Standard genotoxicity test methods and endpoints

Test method	Genotoxic endpoints	EU method / OECD guideline
<i>In vitro</i> test methods		
Bacterial reverse mutation test - Ames	Mutagenicity: gene mutations	EU B.12/13 OECD 471
<i>In vitro</i> mammalian cell gene mutation test – <i>HPRT</i> , <i>TK</i> and <i>XPRT</i> tests	Mutagenicity: gene mutations and structural chromosome aberrations (only in the <i>TK</i> test)	EU B.17 OECD 476*
<i>In vitro</i> mammalian chromosome aberration test	Mutagenicity: structural chromosome aberrations	EU B.10 OECD 473**
<i>In vitro</i> micronucleus test	Mutagenicity: structural and numerical chromosome aberrations	EU B.49 OECD 487 **
<i>In vivo</i> test methods		
<i>In vivo</i> mammalian bone marrow chromosome aberration test	Mutagenicity: structural chromosome aberrations	EU B.11 OECD 475**
<i>In vivo</i> mammalian erythrocyte micronucleus test	Mutagenicity: structural and numerical chromosome aberrations	EU B.12 OECD 474**
Unscheduled DNA synthesis (UDS) test in mammalian liver cells <i>in vivo</i>	Genotoxicity: DNA repair	EU B.39 OECD 486
<i>In vivo</i> mammalian Comet assay	Genotoxicity: DNA strand breaks	EU (none) Draft OECD**
Transgenic rodent somatic and germ cell gene mutation assay	Mutagenicity: gene mutations	EU (none) OECD 488
Mouse heritable translocation assay	Germ cell mutagenicity: structural and numerical chromosome aberrations	EU B.25 OECD 485
Mammalian spermatogonial chromosome aberration test	Germ cell mutagenicity: structural chromosome aberrations	EU B.23 OECD 483
Rodent dominant lethal test	Germ cell mutagenicity: structural and numerical chromosome aberrations	EU B.22 OECD 478
* Two distinct guidelines are in preparation: one for <i>hprt</i> locus mutations detection and a second for <i>tk</i> locus. ** Draft updated Test Guidelines approved.		

8.8 References

- Aardema MJ, Barnett B, Khambatta ZS, Reisinger K, Ouedrago-Arras G, Faquet B, Ginestet A, Mun GC, Dahl, EL, Hewitt NJ, Corvi R & Curren RC (2010). International pre-validation studies of the EpiDerm TM 3D human reconstructed skin micronucleus (RSMN) assay: Transferability and reproducibility. *Mutation Research* 701: 123-131.
- Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfühler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85 (5): 367-485.
- Benigni R, Bossa C, Jeliaskova N, Netzeva N & Andrew Worth (2008). The Benigni /Bossa rulebase for mutagenicity and carcinogenicity –a module of Toxtree. JRC Report, EUR 23241 EN – 2008. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/EUR_23241_EN.pdf
- Billinton N, Bruce S, Hansen JR, Hastwell PW, Jagger C, McComb C, Klug ML, Pant K, Rabinowitz A, Rees R, Tate M, Vinggaard AM & Walmsley RM (2010). A pre-validation transferability study of the GreenScreen HC GADD45a-GFP assay with a metabolic activation system (S9). *Mutation Research* 700: 44-50.
- Birrell L, Cahill P, Hughes C, Tate M & Walmsley RM (2010). GADD45a-GFP GreenScreen HC assay results for the ECVAM recommended lists of genotoxic and non-genotoxic chemicals for assessment of new genotoxicity tests. *Mutation Research* 695: 87-95.
- Bowen DE, Whitwell JW, Lillfordm L, Henderson D, Kidd D, Mc Garry S, Pearce G, Beevers C & Kirkland DJ (2011). Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test, *Mutation Research* 722 (1): 7-19.
- COM, 2011. Guidance on a Strategy for Testing of Chemicals for Mutagenicity. Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM). Department of Health, London. <http://www.iacom.org.uk/guidstate/documents/COMGuidanceFINAL2.pdf>
- Curren RD, Mun GC, Gibson DP & Aardema MJ (2006). Development of a method for assessing micronucleus induction in a 3D human skin model EpiDermTM. *Mutation Research* 607: 192-204.
- De Flora S & Izzotti A (2007). Mutagenesis and cardiovascular diseases Molecular mechanisms, risk factors, and protective factors. *Mutation Research* 621 (1-2): 5-17.
- Dearfield KL, Thybaud V, Cimino MC, Custer L, Czich A, Harvey JS, Hester S, Kim JH, Kirkland D, Levy DD, Lorge E, Moore MM, Ouedrago-Arras G, Schuler M, Suter W, Sweder K, Tarlo K, van Benthem J, van Goethem F & Witt KL (2011). Follow-up actions from positive results of in vitro genetic toxicity testing. *Environmental Molecular Mutagenicity* 52 (3): 177-204.
- DiPaolo JA, Donovan P & Nelson R (1969). Quantitative studies of in vitro transformation by chemical carcinogens. *Journal of National Cancer Institute* 42: 867-874.

Doktorova TY, Pauwels M, Vinken M, Vanhaecke T & Rogiers V (2012). Opportunities for an alternative integrating testing strategy for carcinogen hazard assessment? *Critical Reviews in Toxicology* 42 (2): 91-106.

Doktorova TY, Yildirimman R, Vinken M, Vilardell M, Vanhaecke T, Gmuender H, Bort R, Brolen G, Holmgren G, Li R, Chesne C, van Delft J, Kleinjans J, Castell J, Bjorquist P, Herwig R & Rogiers V (2013). Transcriptomic responses generated by hepatocarcinogens in a battery of liver-based in vitro models. *Carcinogenesis* 34(6):1393-1402.

EC (2006). Regulation No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official Journal of the European Communities* L396, 1-849.

EC (2008). Regulation No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. *Official Journal of the European Union* L353.

EC (2009a). Regulation No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. *Official Journal of the European Communities* L342, 59-209.

EC (2009b). Regulation No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. *Official Journal of the European Union* L309, 1-47.

ECHA (2012). Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific Guidance Version 2.0 November 2012, R.7.7. Mutagenicity and Carcinogenicity. http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf

EFSA (2011). Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. *EFSA Journal* 9 (9): 2379.

Ellinger-Ziegelbauer H, Aubrecht J, Kleinjans JC & Ahr HJ (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. *Toxicology Letters* 186: 36-44.

Erickson RP (2010). Somatic gene mutation and human disease other than cancer: an update. *Mutation Research* 705: 96-106

EU (2012). Regulation No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. *Official Journal of the European Union* L 167, 1-116.

EU (2013) Regulation No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market (Text with EEA relevance).

EU (2013) Regulation No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.

EURL ECVAM (2012). Recommendation concerning the cell transformation assays using Syrian hamster embryo cells (SHE) and the BALB/c 3T3 mouse fibroblast cell line for in vitro carcinogenicity testing. Annex I: ESAC opinion on the ESAC peer review of an ECVAM-coordinated prevalidation study concerning three protocols of the cell transformation assay (CTA) for in vitro carcinogenicity testing. http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing

EURL ECVAM (2013). EURL ECVAM strategy to avoid and reduce animal use in genotoxicity testing. http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-strategy-papers

Fowler P, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael P (2012a). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. *Mutation Research* 742 (1-2): 11-25.

Fowler P, Smith R, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael . (2012b). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. *Mutation Research* 747 (1):104-17.

Frank SA (2010). Evolution in health and medicine Sackler colloquium: Somatic evolutionary genomics: mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proceedings of the National Academy of Science USA* 107 Suppl 1: 1725-1730.

Greywe D, Kreutz J, Banduhn N, Krauledat M, Scheel J, Schroeder KR, Wolf T & Reisinger K (2012). Applicability and robustness of the hen's egg test for analysis of micronucleus induction (HET-MN): results from an inter-laboratory trial. *Mutation Research* 747 (1): 118-34.

Hastwell PW, Webster TW, Tate M, Billinton N, Lynch AM, Harvey JS, Rees RW & Walmsley RM (2009). Analysis of 75 marketed pharmaceuticals using the GADD45a-GFP 'GreenScreen HC' genotoxicity assay. *Mutagenesis* 24: 455-463.

Hastwell PW, Chai LL, Roberts KJ, Webster TW, Harvey JS, Rees RW & Walmsley RM (2006). High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC GADD45a-GFP genotoxicity assay. *Mutation Research* 607: 160-175.

Helma C (2006). Lazy structure-activity relationships (lazar) for the prediction of rodent carcinogenicity and Salmonella mutagenicity. *Molecular Diversity* 10:147-158.

Hoeijmakers JH (2009). DNA damage, aging and cancer. *New England Journal of Medicine* 361: 1475-1485.

Hu T, Khambatta ZS, Hayden PJ, Bolmarcich J, Binder RL, Robinson MK, Carr GJ, Tiesman JP, Jarrold BB, Osborne R, Reichling TD, Nemeth ST & Aardema MJ (2010). Xenobiotic metabolism gene expression in the EpiDerm™ in vitro 3D human epidermis model compared to human skin. *Toxicology in Vitro* 24: 1450-1463.

Hughes C, Rabinowitz A, Tate M, Birrell L, Allsup J, Billinton N & Walmsley RM (2012) Development of a high-throughput Gaussia luciferase Reporter Assay for the Activation of the GADD45a Gene by Mutagens, Promutagens, Clastogens, and Aneugens. *Journal of Biomolecular Screening* 17:1302-1316.

ICH (2011). S2(R1) Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals intended for Human Use. Approval by the Steering Committee of S2(R1) under Step 4 and recommendation for adoption to the three ICH regulatory bodies (9 November 2011).

Isfort RJ, Kerckaert GA & LeBoeuf RA (1996). Comparison of the standard and reduced pH Syrian hamster embryo (SHE) cell in vitro transformation assays in predicting the carcinogenic potential of chemicals. *Mutation Research* 356: 11-16.

Jagger C, Tate M, Cahill PA, Hughes C, Knight AW, Billinton N & Walmsley RM (2009). Assessment of the genotoxicity of S9-generated metabolites using the GreenScreen HC GADD45a-GFP assay. *Mutagenesis* 24: 35-50.

Jennings P, Limonciel A, Felice L & Leonard MO (2013). An overview of transcriptional regulation in response to toxicological insult. *Archives of Toxicology* 87: 49-72.

Kirkland D, Aardema M, Henderson L & Müller L (2005). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutation Research* 584: 1-256.

Kirkland D & Fowler P (2010). Further analysis of Ames-negative rodent carcinogens that are only genotoxic in mammalian cells in vitro at concentrations exceeding 1 mM, including retesting of compounds of concern. *Mutagenesis* 25 (6): 539-553.

Kirkland D, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier J-R, Müller L, Nohynek, G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P & White P (2007). How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. *Mutation Research* 628: 31-55.

Kirkland D, Reeve L, Gatehouse D & Vanparys P (2011). A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins. *Mutation Research* 721: 27-73.

Kirkland D & Speit G (2008). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo. *Mutation Research* 654: 114-132.

Kirsch-Volders M, Decordier I, Elhajouji A, Plas G, Aardema MJ & Fenech M (2011). In vitro genotoxicity testing using the micronucleus assay in cell lines, human lymphocytes and 3D human skin models. *Mutagenesis* 26(1): 177-84.

LeBoeuf RA, Kerckaert KA, Aardema MJ & Isfort RJ (1999). Use of Syrian hamster embryo and BALB/c 3T3 cell transformation for assessing the carcinogenic potential of chemicals. IARC Science Publications, 146: 409-25.

Lynch AM, Sasaki JC, Elespuru R, Jacobson-Kram D, Thybaud V, De Boeck M, Aardema MJ, Aubrecht J, Benz RD, Dertinger SD, Douglas GR, White PA, Escobar PA, Fornace A Jr, Honma M, Naven RT, Rusling JF, Schiestl RH, Walmsley RM, Yamamura E, van Benthem J & Kim JH (2011). New and emerging technologies for genetic toxicity testing. *Environmental Molecular Mutagenicity* 52 (3): 205-23.

Matthews EJ, Spalding JW & Tennant RW (1993). Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays. *Environmental Health Perspectives* 101 Suppl 2: 347-482.

Maurici D, Aardema M, Corvi R, Kleber M, Krul C, Laurent C, Loprieno N, Pasanen M, Pfuhler S, Phillips B, Sabbioni E, Sanner T & Vanparys P (2005). Genotoxicity and mutagenicity. *Alternatives to Laboratory Animals* 33 Suppl 1: 117-130.

Mun GC, Aardema MJ, Hu T, Barnett B, Kaluzhny Y, Karetsky V, Dahl EL & Curren RD (2009). Further development of the EpiDerm 3D reconstructed human skin micronucleus (RSMN) assay. *Mutation research* 673: 92-99.

OECD (2010d) OECD Principles for the validation, for regulatory purposes, of (quantitative) structure activity relationship models. Available from: <http://www.oecd.org/dataoecd/33/37/37849783.pdf>

OECD, Draft Introduction document to the OECD guidelines on genetic toxicology testing and Guidance document on the selection and application of assays, in support of the Genotoxicity Test Guidelines (In preparation).

Parry JM, Parry E, Phrakonkham P & Corvi R (2010). Analysis of published data for top concentration considerations in mammalian cell genotoxicity testing. *Mutagenesis* 25 (6): 531-538.

Pfuhler S, Kirkland D, Kasper P, Hayashi M, Vanparys P, Carmichael P, Dertinger S, Eastmond D, Elhajouji A, Krul C, Rothfuss A, Schoening G, Smith A, Speit G, Thomas C, van Benthem J & Corvi R (2009). Reduction of use of animals in regulatory genotoxicity testing: Identification and implementation opportunities-Report from an ECVAM workshop. *Mutation Research* 680: 31-42.

Pfuhler S, Kirst A, Aardema M, Banduhn N, Goebel C, Araki D, Costabel-Farkas M, Dufour E, Fautz R, Harvey J, Hewitt NJ, Hibatallah J, Carmichael P, Macfarlane M, Reisinger K, Rowland J, Schellauf F, Schepky A & Scheel J (2010). A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: Genotoxicity. A COLIPA analysis. *Regulatory Toxicology & Pharmacology* 57: 315-324.

Reus AA, Reisinger K, Downs TR, Carr G, Zeller A, Corvi R, Krul CAM & Pfuhler S (2013). Comet Assay in reconstructed 3D human epidermal skin models – investigation of intra- and inter-laboratory reproducibility with coded chemicals. *Mutagenesis* 28 (6): 709-720.

Rothfuss A, O'Donovan M, De Boeck M, Brault D, Czich A, Custer L, Hamada S, Plappert-Helbig U, Hayashi M, Howe J, Kraynak A, van der Leede B, Nakajima M, Priestley C, Thybaud V, Saigo K, Sawant S, Shi J, Storer R, Struwe M, Vock E & Galloway S (2010). Collaborative study on 15 compounds in the rat liver Comet Assay integrated into 2- and 4-week repeat-dose studies. *Mutation Research* 702: 40-69.

Rothfuss A., Honma M, Czich A, Aardema MJ, Burlinson B, Galloway, Hamada S, Kirkland D, Heflich RH, Howe J, Nakajima M, O'Donovan M, Plappert-Helbig U, Priestley C, Recio L, Schuler M, Uno Y & Martus HJ (2011). Improvement of in vivo genotoxicity assessment: Combination of acute tests and integration into standard toxicity testing. *Mutation Research* 723 108-120.

SCCS (2012). The SCCS's Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation. 8th Revision. SCCS/1501/12. European Commission. Directorate-General for Health and Consumers. Scientific Committee on Consumer Safety. Available from: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_006.pdf

Serafimova R, Fuat Gatnik M & Worth A (2010). Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC report EUR 24427 EN. Available from: <http://publications.jrc.ec.europa.eu/repository>

Slatter MA & Gennery AR (2010). Primary immunodeficiencies associated with DNA-repair disorders. *Expert. Rev. Mol. Med* 12:e9.

Vanparys P, Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2012). Application of in vitro cell transformation assays in regulatory toxicology for pharmaceuticals, chemical, food products and cosmetics. *Mutation Research* 744 (1): 111-116.

VICH GL23 (R) (Safety) - Genotoxicity Studies to evaluate the safety of residues of veterinary drugs in human food: Genotoxicity Testing. October 2012 Revision at Step 9 For consultation at Step 4, (Chapters 2-3), 1-7.

Waters MD, Jackson M & Lea I (2010). Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods. *Mutation Research* 705: 184-200.

Westerink WM, Stevenson JC, Horbach GJ & Schoonen WG (2010). The development of RAD51C, Cystatin A, p53 and Nrf2 luciferase-reporter assays in metabolically competent HepG2 cells for the assessment of mechanism-based genotoxicity and of oxidative stress in the early research phase of drug development. *Mutation Research* 696: 21-40.

Wolf T, Niehaus-Rolf C, Banduhn N, Eschrich D, Scheel J & Luepke NP (2008). The hen's egg test for micronucleus induction (HET-MN): novel analyses with a series of well-characterized substances support the further evaluation of the test system. *Mutation Research* 650:150-164.

Worth A, Lapenna S, Lo Piparo E, Mostrag-Szlichtyng A & Serafimova R (2010). The applicability of software tools for genotoxicity and carcinogenicity prediction: case studies relevant to the assessment of pesticides. JRC report EUR 24640 EN. Available from: <http://publications.jrc.ec.europa.eu/repository>

Worth A, Lapenna S & Serafimova R (2012). QSAR and metabolic simulation tools in the assessment of genotoxicity. Chapter 6 in: *Computational Toxicology, Volume II. Methods in Molecular Biology, Volume 930*. B Reisfeld & AN Mayeno (Eds). Humana Press, New York. pp 125-162.

Chapter 8. Table of References with Notes

Traditional <i>in vitro</i> and <i>in vivo</i> tests – OECD TGs	
OECD (1997), <i>Test No. 471: Bacterial Reverse Mutation Test</i> , OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071247-en	The <u>bacterial reverse mutation</u> test uses amino-acid requiring at least five strains of Salmonella typhimurium and Escherichia coli to detect point mutations by base substitutions or frame shifts. The principle of this bacterial reverse mutation test is that it detects mutations which, revert mutations present in the test strains and restore the functional capability of the bacteria to synthesize an essential amino acid.
*OECD (1997), <i>Test No. 473: In vitro Mammalian Chromosome Aberration Test</i> , OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071261-en	The purpose of the <i>in vitro</i> chromosome aberration test is to identify agents that cause structural chromosome aberrations in cultured mammalian somatic cells.
OECD (1997), <i>Test No. 476: In vitro Mammalian Cell Gene Mutation Test</i> , OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071322-en	The <i>in vitro</i> mammalian cell gene mutation test can be used to detect gene mutations induced by chemical substances. In the cell lines the most commonly-used genetic endpoints measure mutation at thymidine kinase (TK) and hypoxanthine-guanine phosphoribosyl transferase (HPRT), and a transgene of xanthine-guanine phosphoribosyl transferase (XPRT). The TK, HPRT and XPRT mutation tests detect different spectra of genetic events. Two separate OECD TGs are being drafted: a) concerning TK mutation; b) concerning HPRT/XPRT mutation tests.
*OECD (2010), <i>Test No. 487: In Vitro Mammalian Cell Micronucleus Test</i> , OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264091016-en	The <i>in vitro</i> micronucleus test is used for the detection of micronuclei in the cytoplasm of interphase cells. The assay detects the activity of clastogenic and aneugenic test substances in cells that have undergone cell division during or after exposure to the test substance.
*OECD (1997), <i>Test No. 474: Mammalian Erythrocyte Micronucleus Test</i> , OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071285-en	The <i>in vivo</i> micronucleus test is used for the detection of micronuclei in the cytoplasm of interphase cells. The assay detects the activity of clastogenic and aneugenic test substances in cells that have undergone cell division during or after exposure to the test substance. The test is used for the detection of micronuclei induced in erythroblasts, by analysis of erythrocytes as sampled in bone marrow and/or peripheral blood cells of animals, usually rodents (mice or rats).
*OECD (1997), <i>Test No. 475: Mammalian Bone Marrow</i>	The mammalian <i>in vivo</i> chromosome aberration test is used for the detection of structural

<p><i>Chromosome Aberration Test</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071308-en</p>	<p>chromosome aberrations induced by test compounds in bone marrow cells of animals, usually rodents (rats, mice and Chinese hamsters).</p>
<p>OECD (1984), <i>Test No. 478: Genetic Toxicology: Rodent Dominant Lethal Test</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071360-en</p>	<p>Dominant lethal (DL) effects cause embryonic or foetal death. Induction of a dominant lethal event after exposure to a test substance indicates that the substance has affected germinal tissue of the test species. Dominant lethals are generally accepted to be the result of chromosomal aberrations (structural and/or numerical anomalies), but gene mutations and toxic effects cannot be excluded.</p>
<p>OECD (1997), <i>Test No. 483: Mammalian Spermatogonial Chromosome Aberration Test</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071469-en</p>	<p>This test measures chromosome events in spermatogonial germ cells and is, therefore, expected to be predictive of induction of inheritable mutations in germ cells.</p>
<p>OECD (1986), <i>Test No. 485: Genetic toxicology, Mouse Heritable Translocation Assay</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071506-en</p>	<p>The mouse heritable translocation test detects structural and numerical chromosome changes in mammalian germ cells as recovered in first generation progeny.</p>
<p>OECD (1997), <i>Test No. 486: Unscheduled DNA Synthesis (UDS) Test with Mammalian Liver Cells in vivo</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071520-en</p>	<p>The purpose of the unscheduled DNA synthesis (UDS) test with mammalian liver cells <i>in vivo</i> is to identify substances that induce DNA repair after excision and removal of a stretch of DNA containing a region of damage induced by chemical substances (solid or liquid) or physical agents in the liver.</p>
<p>OECD (2011), <i>Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays</i>, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264122819-en</p>	<p>This <i>in vivo</i> assay detects chemicals that induce gene mutations. In this assay, transgenic rats or mice that contain multiple copies of chromosomally integrated plasmid or phage shuttle vectors are used. The transgenes contain reporter genes for the detection of various types of mutations induced by test substances.</p>
<p>OECD (2013) DRAFT TG Publication: In vivo mammalian Alkline Comet Assay. http://www.oecd.org/env/ehs/testing/OECD_draft_TG_Comet_December2013.pdf</p>	<p>Rodent <i>in vivo</i> Comet assays are especially relevant to assessing DNA-damaging potential of chemicals for regulatory purposes.</p>
<p>OECD (2013) DRAFT TG Publication: In Vitro</p>	<p><i>In vitro</i> cell transformation refers to the induction of phenotypic alterations in cultured cells that are</p>

<p>Carcinogenicity: Syrian Hamster Embryo (SHE) Cell Transformation Assay. http://www.oecd.org/env/ehs/testing/CTA TG_Feb2013.pdf</p>	<p>characteristic of tumorigenic cells. This Test Guideline (TG) provides an <i>in vitro</i> procedure of the cell transformation assay in Syrian Hamster Embryonic stem cells (SHE), which may be used for hazard identification of chemical carcinogens.</p>
<p>Regulatory requirements across sectors</p>	
<p>ICH (2011) S2(R1) Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use . Approval by the Steering Committee of S2(R1) under Step 4 and recommendation for adoption to the three ICH regulatory bodies (9 November 2011)</p>	<p><i>Pharmaceuticals:</i> <u>Chapter 2. The standard test battery for genotoxicity:</u> Description of the Two Options for the Standard Battery. The following two options for the standard battery are considered equally suitable: Option 1 i. A test for gene mutation in bacteria. ii. A cytogenetic test for chromosomal damage (the <i>in vitro</i> metaphase chromosome aberration test or <i>in vitro</i> micronucleus test), or an <i>in vitro</i> mouse lymphoma Tk gene mutation assay. iii. An <i>in vivo</i> test for genotoxicity, generally a test for chromosomal damage using rodent hematopoietic cells, either for micronuclei or for chromosomal aberrations in metaphase cells. Option 2 i. A test for gene mutation in bacteria. ii. An <i>in vivo</i> assessment of genotoxicity with two different tissues, usually an assay for micronuclei using rodent hematopoietic cells and a second <i>in vivo</i> assay. Typically this would be a DNA strand breakage assay in liver, unless otherwise justified.</p>
<p>VICH (2012) GL23(R) (Safety) Genotoxicity Studies to evaluate the safety of residues of veterinary drugs in human food: Genotoxicity Testing. October 2012 Revision at Step 9For consultation at Step 4, (Chapters 2-3), pp 1-7 (under revision).</p>	<p><i>Veterinary Drugs:</i> A battery of three tests is recommended for use as a screen of veterinary drugs for genotoxicity:</p> <ul style="list-style-type: none"> • A test for gene mutation in bacteria. • A cytogenetic test for chromosomal damage (the <i>in vitro</i> metaphase chromosome aberration test or <i>in vitro</i> micronucleus test), or an <i>in vitro</i> mouse lymphoma tk gene mutation assay. • An <i>in vivo</i> test for chromosomal effects using rodent haematopoietic cells. <p>Modifications can be applied and modified protocols should be used where it is evident that standard conditions will give a false negative result.</p>
<p>European Commission, 2006. Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December</p>	<p><i>Chemicals:</i> Information requirements depend on the production volume of the chemical concerned:</p>

<p>2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Communities L396, 1-849.</p>	<ul style="list-style-type: none"> • Annex VII (1-10 tpa): <i>In vitro</i> gene mutation study in bacteria. Further mutagenicity studies shall be considered in case of a positive result. • Annex VIII (10-100 tpa) <ol style="list-style-type: none"> 1) <i>In vitro</i> cytogenicity study in mammalian cells or <i>in vitro</i> micronucleus study. The study does not usually need to be conducted – if adequate data from an <i>in vivo</i> cytogenicity test are available or – the substance is known to be carcinogenic category 1 or 2 or mutagenic category 1, 2 or 3 (referring to GHS category 1A, 1B and 2, respectively). 2) <i>In vitro</i> gene mutation study in mammalian cells, if a negative result in Annex VII and Annex VIII. The study does not usually need to be conducted if adequate data from a reliable <i>in vivo</i> mammalian gene mutation test are available. Appropriate <i>in vivo</i> mutagenicity studies shall be considered in case of a positive result in any of the genotoxicity studies in Annex VII or VIII. • Annex IX (100-1000 tpa) If there is a positive result in any of the <i>in vitro</i> genotoxicity studies in Annex VII or VIII and there are no results available from an <i>in vivo</i> study already, an appropriate <i>in vivo</i> somatic cell genotoxicity study shall be proposed by the registrant. If there is a positive result from an <i>in vivo</i> somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered. • Annex X (1000-more tpa) If there is a positive result in any of the <i>in vitro</i> genotoxicity studies in Annexes VII or VIII, a second <i>in vivo</i> somatic cell test may be necessary, depending on the quality and relevance of all the available data. If there is a positive result from an <i>in vivo</i> somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.
<p>European Commission (2009a). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the</p>	<p><i>Cosmetics:</i> <u>On July 11, 2013</u>, the Regulation no. 1223/2009 fully replaced the “<i>Cosmetics Directive</i>” 76/68/EEC and following Amendments (Directive 93/35/EEC, Directive 2003/15/EC. This</p>

<p>European Communities L342, 59-209.</p> <p>And</p> <p>SCCS (2012) SCCS'S Notes of Guidance for the testing of cosmetics substances and their safety evaluation, 8th Revision, SCCS/1501/12</p>	<p>Regulation establishes rules to be complied with by any cosmetic product made available on the market, in order to ensure the functioning of the internal market and a high level of protection of human health.</p> <p>The Regulation also prohibits the placing on the European Union market of [Chapter V, Article 18, Animal testing]:</p> <ul style="list-style-type: none"> • Products where the final formulation has been the subject of animal testing; • Products containing ingredients or combinations of ingredients which have been the subject of animal testing. <p>Based on the recent SCCS's Notes of Guidance (SCCS/1501/12), which is currently under revision, three assays, for the basic level testing of cosmetic substances are recommended:1. <u>Tests for gene mutation:</u></p> <p>i) Bacterial Reverse Mutation Test ii) <i>In Vitro</i> Mammalian Cell Gene Mutation Test</p> <p>2. <u>Tests for clastogenicity and aneugenicity</u></p> <p>i) <i>In Vitro</i> Micronucleus Test OR ii) <i>In vitro</i> Mammalian Chromosome Aberration Test</p> <ul style="list-style-type: none"> • In cases where negative results are seen in the conducted tests, a mutagenic potential is excluded. • Likewise, in cases where a positive result is seen in one of the tests, the compound has to be considered as a (<i>in vitro</i>/intrinsic) mutagen.
<p>European Commission (2009b). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-47.</p> <p>And</p> <p>Commission Regulation (EU), 2013, No 283/2013 of 1 March 2013 setting out the data requirements for active</p>	<p><i>Plant Protection Products:</i></p> <p>This Regulation establishes rules to be complied with by any plant protection product made available on the market.</p> <p>An active substance, safener or synergist shall only be approved if, on the basis of assessment of higher tier genotoxicity testing carried out in accordance with the data requirements for the active substances, safeners or synergists and other available data and information, including a review of the scientific literature, reviewed by the Authority, it is not or has not to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as mutagen category 1A or 1B.</p> <p>New and revised genotoxicity testing requirements are reported in detail in Paragraph 5.4 (EU No.</p>

<p>substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.</p> <p>Commission Regulation (EU), 2013, No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market (Text with EEA relevance)</p>	<p>283/2013)</p>
<p>European Union Regulation (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.</p>	<p><i>Biocides</i>: Annex II (par. 8.5-8.6), Data Requirement</p> <p><u>Mutagenicity</u></p> <p>The assessment of this endpoint shall comprise the following consecutive steps: — an assessment of the available <i>in vivo</i> genotoxicity data — an <i>in vitro</i> test for gene mutations in bacteria, an <i>in vitro</i> cytogenicity test in mammalian cells and an <i>in vitro</i> gene mutation test in mammalian cells are required — appropriate <i>in vivo</i> genotoxicity studies shall be considered in case of a positive result in any of the <i>in vitro</i> genotoxicity studies.</p> <p><u>In vivo genotoxicity study</u></p> <p>The assessment of this endpoint shall comprise the following consecutive steps: — If there is a positive result in any of the <i>in vitro</i> genotoxicity studies and there are no results available from an <i>in vivo</i> study already, an appropriate <i>in vivo</i> somatic cell genotoxicity study shall be proposed/conducted by the applicant — If either of the <i>in vitro</i> gene mutation tests is positive, an <i>in vivo</i> test to investigate unscheduled DNA synthesis shall be conducted — A second <i>in vivo</i> somatic cell test may be necessary, depending on the results, quality and relevance of all the available data — If there is a positive result from an <i>in vivo</i> somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence to demonstrate that the substance reached the tested organ. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.</p> <p><u>Additional data set</u></p> <p>The study/ies do(es) not generally need to be conducted if: — the results are negative for the three <i>in vitro</i> tests and if no metabolites of concern are formed in mammals or — valid <i>in vivo</i> micronucleus</p>

	data is generated within a repeat dose study and the <i>in vivo</i> micronucleus test is the appropriate test to be conducted to address this information requirement — the substance is known to be carcinogenic category 1A or 1B or mutagenic category 1A, 1B or 2.
<p>European Commission (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353.</p> <p>UN (2013) Globally Harmonized System of Classification and Labelling of Chemicals (GHS), 5th revised edition. United Nations, New York and Geneva, 2013 http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf</p> <p>[Chapter 3.5, fig. 3.5.1].</p>	<p>EU CLP Regulation</p> <p>Par. 3.5.1 Definitions and general conditions: '<i>mutagenic, mutagen; genotoxic, genotoxicity</i>'</p> <p>Par. 3.5.2 Classification criteria for substances</p> <p>Table. 3.5.1 Hazard categories for germ cell mutagens: '<i>Category 1A and 1B and Category 2</i>' -CLP and Globally harmonized System (GHS) classification</p>
Guidance Documents and Reviews on Genotoxicity	
<p>ECHA (2012). Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific Guidance Version 2.0 November 2012, R.7.7. Mutagenicity and Carcinogenicity. http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf</p>	<p>This document describes the information requirements under REACH with regard to substance properties, exposure, uses and risk management measures, and the chemical safety assessment. In this chapter, specific guidance on meeting the information requirements set out in Annexes VI to XI to the REACH Regulation is provided.</p>
<p>EFSA (2011). Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011; 9 (9): 2379 (69pp)</p>	<p>The Scientific Committee reviewed the current state-of-the-science on genotoxicity testing and provided a commentary and recommendations on genotoxicity testing strategies. A step-wise approach is recommended for the generation and evaluation of data on genotoxic potential, beginning with a basic battery of <i>in vitro</i> tests, comprising a bacterial reverse mutation assay and an <i>in vitro</i> micronucleus assay. In the event of negative <i>in vitro</i> results, it can be concluded that the substance has no genotoxic potential. In case of inconclusive, contradictory or equivocal results, it</p>

	<p>may be appropriate to conduct further testing <i>in vitro</i>. In case of positive <i>in vitro</i> results, review of the available relevant data on the test substance and, where necessary, an appropriate <i>in vivo</i> study to assess whether the genotoxic potential observed <i>in vitro</i> is expressed <i>in vivo</i> is recommended. The approach to <i>in vivo</i> testing should be also step-wise. The combination of assessing different endpoints in different tissues in the same animal <i>in vivo</i> should also be considered.</p>
<p>COM (2011). Guidance on a Strategy for Testing of Chemicals for Mutagenicity. Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM). Department of Health, London. http://www.iacom.org.uk/guidstate/documents/COMGuidanceFINAL2.pdf</p>	<p>The COM guidance provides a strategy for testing chemical substances where no genotoxicity data are available.</p> <p><u>Stage 0: Preliminary considerations prior to testing.</u> Analysis of Physico-chemical and Toxicological Properties; (Q)SAR models for prediction of mutagenic activity; Screening tools i.e. <i>in silico</i> approach.</p> <p><u>Stage 1: <i>In vitro</i> genotoxicity testing.</u> The strategy includes using appropriate tests to gain an insight into the nature of the genotoxic effects of a test substance and also to <i>avoid misleading positive results</i>. It comprises a two-test core system (namely an Ames test and <i>in vitro</i> micronucleus test, MNvit) with the objective of assessing mutagenic potential by investigating three different endpoints (gene mutation, structural chromosomal damage and changes in chromosome number).</p> <p><u>Stage 2: <i>In vivo</i> genotoxicity testing.</u> The <i>in vivo</i> genotoxicity testing strategy has to be designed on a case-by case basis and can be used to address aspects of <i>in vivo</i> mutagenicity.</p>
<p>EURL ECVAM (2013). EURL ECVAM strategy to avoid and reduce animal use in genotoxicity testing. http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-strategy-papers</p>	<p>Although several <i>in vitro</i> tests are available at different stages of development and acceptance, they cannot at present be considered to fully replace animal tests needed to evaluate the safety of substances. Based on an analysis of regulatory requirements for this endpoint within different pieces of EU legislation, EURL ECVAM proposes a pragmatic approach to improve the traditional genotoxicity testing paradigm that offers solutions in both the short- and medium-term and that draws on the considerable experience of 40 years of regulatory toxicology testing in this area. EURL ECVAM considers that efforts should be directed towards the overall improvement of the current testing strategy for better hazard and risk assessment approaches, which either avoids or minimises the use of animals, whilst satisfying regulatory information requirements, irrespective of regulatory context. Several opportunities for the improvement of the testing strategy have been identified which aim to i) enhance the performance of the <i>in vitro</i> testing battery so that fewer <i>in vivo</i> follow-up tests are necessary and ii) guide more intelligent <i>in vivo</i> follow-up testing to reduce unnecessary use of animals. The implementation of this strategic plan will rely on the cooperation of EURL ECVAM with other existing initiatives and the coordinated contribution from various stakeholders.</p>
<p>Maurici D, Aardema M, Corvi R, Kleber M, Krul C, Laurent</p>	<p>This report provided an objective state of play of the status of alternative methods/strategies and the</p>

<p>C., Loprieno N, Pasanen M, Pfuhrer S, Phillips B, Sabbioni E, Sanner T & Vanparys P (2005). Genotoxicity and mutagenicity. Alternatives to Laboratory Animals 33 Suppl 1: 117-130</p>	<p>prospects for their validation and regulatory acceptance so that they could be used for replacing animal tests in the safety assessment of cosmetic products as required by the EU Cosmetics Directive. Regarding the endpoint genotoxicity/mutagenicity, the experts were of the opinion that a total replacement of animal testing was not feasible in the short-term and this would depend, besides the development of <i>in vitro</i> tests on skin models, also on the progress in the fields of toxicokinetics and toxicogenomics.</p>
<p>Adler S, Basketter D, Creton S, et al. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Archives of Toxicology 85 (5): 367-485.</p>	<p>In regards of the 7th amendment to the EU Cosmetics Directive to prohibit animal-tested cosmetics on the market starting 2013, the European Commission invited stakeholder bodies (industry, nongovernmental organisations, EU Member States, and the Commission’s Scientific Committee on Consumer Safety) to identify scientific experts to review the state of art of alternative methods in five toxicological areas, for which the Directive foresees that the 2013 deadline could be further extended in case alternative and validated methods would not be available in time i.e. toxicokinetics, repeated dose toxicity, carcinogenicity, skin sensitisation, and reproductive toxicity.</p> <p>In summary, the experts confirmed that it will take at least another 7–9 years for the replacement of the current <i>in vivo</i> animal tests used for the safety assessment of cosmetic ingredients for skin sensitisation. However, the experts were also of the opinion that alternative methods may be able to give hazard information, i.e. to differentiate between sensitizers and non-sensitizers, ahead of 2017.</p> <p>For toxicokinetics, the timeframe was 5–7 years to develop the models still lacking to predict lung absorption and renal/biliary excretion, and even longer to integrate the methods to fully replace the animal toxicokinetic models. For the systemic toxicological endpoints of repeated dose toxicity, carcinogenicity and reproductive toxicity, the time horizon for full replacement could not be estimated</p>
<p>Misleading positives and follow up of <i>in vitro</i> positives</p>	
<p>Kirkland D, Aardema M, Henderson L & Müller L (2005). Evaluation of the ability of a battery of three <i>in vitro</i> genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. Mutation Research 584: 1-256.</p>	<p>This paper evaluated how the current standard <i>in vitro</i> genotoxicity tests – Ames test, MLA and a test for clastogenicity (<i>in vitro</i> micronucleus or chromosomal aberration test) – performed in their ability to discriminate rodent carcinogens and non carcinogens. The work deals separately with the performance indicators (sensitivity and specificity) of the individual tests and various combinations of two or three tests, and proposes reasons why some rodent carcinogens are not detected in this battery. While the sensitivity of the three test combination was high, its specificity was extremely low, highlights the importance of understanding the mechanism by which genotoxicity may be induced and using weight of evidence approaches to assess the carcinogenic risk from a positive genotoxicity signal. It also highlights deficiencies in the current prediction from and understanding</p>

	of such <i>in vitro</i> results for the <i>in vivo</i> situation.
Kirkland D, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier J-R, Müller L, Nohynek G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P & White P (2007). How to reduce false positive results when undertaking <i>in vitro</i> genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. <i>Mutation Research</i> 628: 31-55.	A workshop organised by ECVAM in 2006 on "How to reduce the false positive results when undertaking <i>in vitro</i> genotoxicity testing and thus avoid unnecessary follow-up animal tests" addressed (i) whether it is possible to choose existing cell systems which give lower rates of false positive results, (ii) whether modifications of existing protocols and cell systems may result in lower false positive results, (iii) the performance of new test systems showing promise of improved specificity and definition of needs for the development of new tests. The recommendations of this workshop paved the way for many international research initiatives.
Dearfield KL, Thybaud V, Cimino MC, Custer L, Czich A, Harvey JS, Hester S, Kim JH, Kirkland D, Levy DD, Lorge E, Moore MM, Ouédraogo-Arras G, Schuler M, Suter W, Sweder K, Tarlo K, van Benthem J, van Goethem F & Witt KL (2011). Follow-up actions from positive results of <i>in vitro</i> genetic toxicity testing. <i>Environmental Molecular Mutagenicity</i> 52 (3): 177-204.	The International Life Sciences Institute (ILSI) Health and Environmental Sciences Institute (HESI) Project Committee on the Relevance and Follow-up of Positive Results in In Vitro Genetic Toxicity (IVGT) Testing developed a decision process flow chart to be applied in case of clear positive results <i>in vitro</i> ; providing for a variety of different possibilities and allowing flexibility in choosing follow-up action(s), depending on the results obtained in the initial battery of assays and available information. The Review Subgroup: <ul style="list-style-type: none"> • Reinforced the concept of weighing the totality of the evidence. • Highlighted the importance of properly analyzing the existing data, and considering potential confounding factors (e.g., possible interactions with the test systems, presence of impurities, irrelevant metabolism), and chemical modes of action when analyzing and interpreting positive results in the <i>in vitro</i> genotoxicity assays and determining appropriate follow-up testing. • Examined the characteristics, strengths, and limitations of each of the existing <i>in vitro</i> and <i>in vivo</i> genotoxicity assays to determine their usefulness in any follow-up testing.
Lynch AM, Sasaki JC, Elespuru R, Jacobson-Kram D, Thybaud V, De Boeck M, Aardema MJ, Aubrecht J, Benz RD, Dertinger SD, Douglas GR, White PA, Escobar PA, Fornace A Jr, Honma M, Naven RT, Rusling JF, Schiestl RH, Walmsley RM, Yamamura E, van Benthem J & Kim JH (2011). New and emerging technologies for genetic toxicity testing. <i>Environmental Molecular Mutagenicity</i> 52 (3): 205-	The Health and Environmental Sciences Institute (HESI) Project Committee on the Relevance and Follow-up of Positive Results in In Vitro Genetic Toxicity (IVGT) convened a workshop in Washington, DC in May 2008 to discuss mature, maturing, and emerging technologies in genetic toxicology. This article collates the abstracts of the New and Emerging Technologies Workshop together with some additional technologies subsequently considered. In particular, consideration was given with regard to follow-up testing of positive results in the standard IVGT tests (i.e., Salmonella Ames test, chromosome aberration assay, and mouse

23.	lymphoma assay) to add weight of evidence and/or provide mechanism of action for improved genetic toxicity risk assessments in humans.
<i>In silico methods</i>	
Serafimova R, Fuat Gatnik M & Worth A (2010). Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC report EUR 24427 EN: http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/publications/	The authors reviewed QSARs for genotoxicity and carcinogenicity, taking into accounts both models available in software tools and models that are published in the literature. They focused also on the potential applicability of diverse models to pesticides as well as to other types of regulated chemicals and pharmaceuticals.
Worth A, Lapenna S, Lo Piparo E, Mostrag-Szlichtyng A & Serafimova R (2010). The applicability of software tools for genotoxicity and carcinogenicity prediction: case studies relevant to the assessment of pesticides. JRC report EUR 24640 EN. Available from: http://publications.jrc.ec.europa.eu/repository	This report presents research results obtained in the framework of a project on the Applicability of Quantitative Structure-Activity Relationship (QSAR) analysis in the evaluation of the toxicological relevance of metabolites and degradates of pesticide active substances.
Helma C (2006). Lazy structure-activity relationships (lazar) for the prediction of rodent carcinogenicity and Salmonella mutagenicity. Molecular Diversity 10:147-158.	The author reports on the Lazar (lazy learning method) where compounds are selected at the time of processing a query compound. The method allows models to be updated as new data become available and includes models for mutagenicity and rodent carcinogenicity.
OECD (2010d) OECD Principles for the validation, for regulatory purposes, of (quantitative) structure activity relationship models http://www.oecd.org/dataoecd/33/37/37849783.pdf	The document describes the principles to facilitate the consideration of a (Q)SAR model for regulatory purposes.
Benigni R, Bossa C, Jeliaskova N, Netzeva N & Worth A (2008). The Benigni /Bossa rulebase for mutagenicity and carcinogenicity –a module of Toxtree. JRC Report, EUR 23241 EN – 2008. http://publications.jrc.ec.europa.eu/repository	This paper describes novel hazard estimation software called Toxtree, capable of making structure-based predictions for a number of toxicological endpoints, developed at the JRC. One of the modules developed as an extension to Toxtree is aimed at the prediction of carcinogenicity and mutagenicity. The main tool is a list of Structural Alerts (SA) for carcinogenicity.
Worth A, Lapenna S & Serafimova R (2012). QSAR and metabolic simulation tools in the assessment of genotoxicity. Chapter 6 in: Computational Toxicology, Volume II. Methods in Molecular Biology, Volume 930. B Reisfeld & AN Mayeno (Eds). Humana Press, New York. pp 125-162.	In this book chapter, a range of computational tools for applying QSAR and grouping/read-across methods are described, and their integrated use in the computational assessment of genotoxicity is illustrated through the application of selected tools to two case-study compounds—2-amino-9H-pyrido[2,3-b]indole (AaC) and 2-aminoacetophenone (2-AAP). The first case study compound (AaC) is an environment pollutant and a food contaminant that can be formed during the cooking of

	<p>protein-rich food. The second case study compound (2-AAP) is a naturally occurring compound in certain foods and also proposed for use as a flavoring agent. The overall aim is to describe and illustrate a possible way of combining different information sources and software tools for genotoxicity and metabolism prediction by means of a simple stepwise approach. The chapter is aimed at researchers and assessors who have a basic knowledge of computational toxicology and some familiarity with the practical use of computational tools. The emphasis is on how to evaluate the data generated by multiple tools, rather than the practical use of any specific tool.</p>
<p><i>In vitro</i> genotoxicity tests in 3D human reconstructed skin models</p>	
<p>Pfuhler S, Kirst A, Aardema M, Banduhn N, Goebel C, Araki D, Costabel-Farkas M, Dufour E, Fautz R, Harvey J, Hewitt NJ, Hibatallah J, Carmichael P, Macfarlane M, Reisinger K, Rowland J, Schellauf F, Schepky A & Scheel J (2010). A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: Genotoxicity. A COLIPA analysis. <i>Regulatory Toxicology Pharmacology</i> 57:315-324.</p>	<p>This article describes common approaches of cosmetic companies, with recommendations for evaluating <i>in vitro</i> genotoxins using non-animal approaches. A weight of evidence approach was employed to set up a decision-tree for the integration of alternative methods into tiered testing strategies.</p>
<p>Hu T, Khambatta ZS, Hayden PJ, Bolmarcich J, Binder RL, Robinson MK, Carr GJ, Tiesman JP, Jarrold BB, Osborne R, Reichling TD, Nemeth ST & Aardema MJ (2010). Xenobiotic metabolism gene expression in the EpiDerm™ <i>in vitro</i> 3D human epidermis model compared to human skin. <i>Toxicol in Vitro</i> 24: 1450-1463.</p>	<p>This paper presents the characterization of the metabolizing capacity of <i>in vitro</i> human skin models. A comparison of the expression of 139 genes encoding xenobiotic metabolizing enzymes in the EpiDerm™ model and human skin is reported. In microarray analysis, the expression of 87% of the genes was consistent between the EpiDerm™ model and human skin indicating the presence of similar metabolic pathways suggesting commonality in function. Analysis of EpiDerm™ models constructed from four donors showed highly comparable expression of xenobiotic metabolizing genes demonstrating reproducibility of the model. Overall, the expression of Phase II enzymes appeared to be more pronounced in human skin and the EpiDerm™ model than that of Phase I enzymes, consistent with the role of skin in detoxification of xenobiotics. Though the basal expression of CYPs in particular was low in EpiDerm™, significant induction of CYP1A1/1B1 activity was observed following treatment with 3-methylcholanthrene. These results indicate that the xenobiotic metabolizing capacity of the EpiDerm™ model appears to be representative of human skin.</p>
<p>Curren RD, Mun GC, Gibson DP & Aardema MJ (2006). Development of a method for assessing micronucleus induction in a 3D human skin model EpiDerm™. <i>Mutation</i></p>	<p>The paper describes the development of a micronucleus assay that uses EpiDerm engineered human reconstructed skin(MatTek Corp., Ashland, MA). Methods for isolating single cells from the 3D skin model and for processing the cells for microscopic analysis of micronuclei (MN) were also</p>

Research 607: 192-204.	described.
Mun GC, Aardema MJ, Hu T, Barnett B, Kaluzhny Y, Karetsky V, Dahl EL & Curren RD (2009). Further development of the EpiDerm 3D reconstructed human skin micronucleus (RSMN) assay. Mutation research 673: 92-99.	The authors developed an <i>in vitro</i> micronucleus assay using a three-dimensional human reconstructed skin model (EpiDerm™), and in this paper presented the refinements applied to the model. They assessed its ability to predict genotoxicity of a battery of chemicals that were previously classified as genotoxins or non-genotoxins based on <i>in vivo</i> rodent skin tests.
Aardema MJ, Barnett B, Khambatta ZS, Reisinger K, Ouedrago-Arras G, Faquet B, Ginestet A, Mun GC, Dahl, EL, Hewitt NJ, Corvi R & Curren RC (2010). International pre-validation studies of the EpiDerm TM 3D human reconstructed skin micronucleus (RSMN) assay: Transferability and reproducibility. Mutation Research 701: 123-131.	This inter-laboratory study aims at establishing the reliability of the EpiDerm TM 3D human reconstructed skin micronucleus (RSMN) assay and at increasing the domain of chemicals tested for predictive capacity.
Reus AA, Reisinger K, Downs TR, Carr G, Zeller A, Corvi R, Krul CAM & Pfuhrer S (2013). Comet Assay in reconstructed 3D human epidermal skin models – investigation of intra- and inter-laboratory reproducibility with coded chemicals. Mutagenesis. 28 (6): 709-20.	This paper describes the development and optimisation of a comet assay protocol in a 3D human EpiDerm™ skin model, as well as intra- and inter-laboratory reproducibility results from the testing of five compounds. A high % tail DNA value of untreated controls was observed. Despite this, valid experiments showed an acceptable reproducibility within and across the laboratories. The work presented will provide a starting point for further investigation on the predictive capacity of comet assays in reconstructed skin models.
Hen's egg test for micronucleus induction (HET-MN)	
Wolf T, Niehaus-Rolf C, Banduhn N, Eschrich D, Scheel J & Luepke NP (2008). The hen's egg test for micronucleus induction (HET-MN): novel analyses with a series of well-characterized substances support the further evaluation of the test system. Mutation Research 650: 150-164.	The hen's egg test for micronucleus induction (HET-MN) combines the use of the commonly accepted genetic endpoint "formation of micronuclei" with the well-characterized and complex model of the incubated hen's egg, which enables metabolic activation, elimination and excretion of xenobiotics - including those that are mutagens or promutagens. The authors present results of experiments involving genotoxic and non-genotoxic model substances. The comparison of these data with previously published data, showed a lack of false negatives or false positives, thus demonstrating a predictivity of genotoxic effects with the above assay.
Cell transformation assay (CTA)	
DiPaolo JA, Donovan P, Nelson R (1969) Quantitative studies of <i>in vitro</i> transformation by chemical carcinogens. Journal of the National Cancer Institute 42: 867-874.	The authors report that chemical carcinogens can induce malignant transformation in mammalian cell systems. However, the precise molecular and cellular alterations that result in transformation have not been identified.

<p>Isfort RJ, Kerckaert GA & LeBoeuf RA (1996). Comparison of the standard and reduced pH Syrian hamster embryo (SHE) cell <i>in vitro</i> transformation assays in predicting the carcinogenic potential of chemicals. <i>Mutation Research</i> 356: 11-63.</p>	<p>The article is a comprehensive review of the Syrian Hamster Embryo (SHE) cell transformation literature; it was performed in order to catalogue the chemical/physical entities which have been evaluated for <i>in vitro</i> cell transformation potential. Both reduced pH (pH 6.7) and standard pH (pH 7.1-7.3) SHE cell testing protocols were considered. The authors suggest that the SHE cell transformation assay is predictive for rodent carcinogenicity under either reduced or standard pH conditions; with the assay displaying better performance in prediction capability under reduced pH conditions.</p>
<p>Matthews EJ, Spalding JW & Tennant RW (1993). Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays. <i>Environmental Health Perspectives</i> 101 Suppl 2: 347-482.</p>	<p>This report describes the activities of 168 chemicals tested in a standard transformation assay using A-31-1-13 BALB/c-3T3 cells. Data analyses revealed that the transformation assay and rodent bioassay had a concordance of 71%; sensitivity for carcinogens of 80.0%; specificity for detecting non carcinogens of 60%.</p> <p>In contrast, Salmonella mutagenicity assays and rodent bioassays had a concordance of 63%, a sensitivity of 58%, and a specificity of 69%. The transformation assay complemented the Salmonella mutagenesis assay in the identification of non-mutagenic carcinogens; thus, the two assays had a combined 83% sensitivity for all carcinogens and a specificity of 75% for non-mutagenic non carcinogens.</p>
<p>LeBoeuf RA, Kerckaert KA, Aardema MJ & Isfort RJ (1999). Use of Syrian hamster embryo and BALB/c 3T3 cell transformation for assessing the carcinogenic potential of chemicals. <i>IARC Science Publications</i> 146: 409-25.</p>	<p>CTAs have been shown to involve a multistage process that closely models key stages of carcinogenesis.</p>
<p>EURL ECVAM (2012) recommendation concerning the cell transformation assays using Syrian hamster embryo cells (SHE) and the BALB/c 3T3 mouse fibroblast cell line for <i>in vitro</i> carcinogenicity testing. Annex I: ESAC opinion on the ESAC peer review of an ECVAM-coordinated prevalidation study concerning three protocols of the cell transformation assay (CTA) for <i>in vitro</i> carcinogenicity testing. http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing</p>	<p>ECVAM recommendation report on the three CTA protocol variants: (a) The SHE CTA performed at pH 6.7 (SHE pH 6.7 CTA); (b) the SHE CTA performed at pH 7.0 (SHE pH 7.0 CTA); and (c) the BALB/c 3T3 CTA. ECVAM coordinated an international study that was designed to address issues of CTA protocols standardisation, transferability and reproducibility in three protocol variants for the SHE CTA (at pH 6.7 and pH 7.0) and the BALB/c 3T3 assay. The study of the three test methods was peer reviewed by the ESAC that issued a peer review report and an ESAC opinion. Based on the available documents (study reports, ESAC opinion, OECD DRP), EURL ECVAM recommended that an OECD TG be developed for the SHE CTA and that further investigations were needed to confirm the performance of the BALB/c 3T3 CTA. Moreover, the recommendation stated that in conjunction with other available data, the CTAs have the potential of partial replacement or reduction when used in a weight of evidence approach for hazard</p>

	identification and risk assessment.
Vanparrys P, Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2012). Application of <i>in vitro</i> cell transformation assays in regulatory toxicology for pharmaceuticals, chemical, food products and cosmetics. Mutation Research 744 (1): 111-116.	The extensive review on CTAs by the OECD (2007) and the proven standardisation, intra- and inter-laboratory reproducibility of the SHE CTAs justify broader use of these methods to assess carcinogenic potential of chemicals. The manuscript describes possible applications of the CTA in relation to different industrial sectors, e.g. food additives, chemicals, cosmetics, pharmaceuticals, etc.
Toxicogenomics-based Tests	
Ellinger-Ziegelbauer H, Aubrecht J, Kleinjans JC & Ahr HJ (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. Toxicology Letters 186: 36-44.	The paper underlies the importance of developing novel scientific approaches bridging genotoxicity and carcinogenicity testing via understanding underlying mechanisms for facilitating cancer risk assessment. The authors consider extremely promising the toxicogenomics approaches as these have the potential of providing generic insight in molecular pathway responses. The report reviews recent progress in the development and application of toxicogenomics to the derivation of genomic biomarkers associated with mechanisms of genotoxicity and carcinogenesis.
Doktorova TY, Pauwels M, Vinken M, Vanhaecke T & Rogiers V (2012). Opportunities for an alternative integrating testing strategy for carcinogen hazard assessment? Critical Reviews in Toxicology 42 (2): 91-106.	In this review paper the major advantages and pitfalls of the existing alternative methodologies to the carcinogenicity bioassay are discussed. Based on the available scientific data in the public domain, the authors propose a "feasible integrated testing strategy" which incorporates some promising new alternatives (toxicogenomics-, transcriptomics-based tests and CTAs), providing at the same time information on the mechanism of action and the toxic nature of the compounds tested.
Doktorova TY, Yildirimman R, Vinken M, Vilardell M, Vanhaecke T, Gmuender H, Bort R, Brolen G, Holmgren G, Li R, Chesne C, van Delft J, Kleinjans J, Castell J, Bjorquist P, Herwig R & Rogiers V (2013). Transcriptomic responses generated by hepatocarcinogens in a battery of liver-based <i>in vitro</i> models. Carcinogenesis 34(6):1393-402.	In the present study the transcriptomics responses following exposure to genotoxic and non-genotoxic hepatocarcinogens and non-carcinogens in five liver-based <i>in vitro</i> models, namely conventional and epigenetically stabilized cultures of primary rat hepatocytes, the human hepatoma-derived cell lines HepaRG and HepG2 and human embryonic stem cell-derived hepatocyte-like cells, are examined.
Jennings P, Limonciel A, Felice L & Leonard MO (2013). An overview of transcriptional regulation in response to toxicological insult. Archives of Toxicology 87: 49-72.	This review aims to consolidate and summarise the major toxicologically relevant transcription factor-governed molecular pathways. The authors focus on the importance of toxicogenomic data sets and information on regulation of stress pathways at the transcriptome level, in order to appreciate the diversity and complexity of biological responses to xenobiotics.

<p>Waters MD, Jackson M & Lea I (2010). Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods. <i>Mutation Research</i> 705: 184-200.</p>	<p>The authors report that the evidence accumulated to date suggests that gene expression profiles reflect underlying modes or mechanisms of action, such that they will be useful in the prediction of chemical carcinogenicity, especially in conjunction with conventional short-term tests for gene mutation, chromosomal aberration and aneuploidy.</p>
<p>Screening approaches</p>	
<p>Hastwell PW, Chai LL, Roberts KJ, Webster TW, Harvey JS, Rees RW & Walmsley RM (2006). High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC GADD45a-GFP genotoxicity assay. <i>Mutation Research</i> 607:160-175.</p>	<p>In light of the relatively high specificity of the Salmonella mutagenicity assay (Ames test) counteracting the low specificity of the established mammalian cell assays, which leads to difficulties in the interpretation of the biological relevance of results, the authors introduce a new high-throughput assay that links the regulation of the human GADD45a gene to the production of Green Fluorescent Protein (GFP). The test is proposed both as a valuable and amenable tool in the selection of candidate compounds for further development, or in product development areas where the use of animals is to be discontinued. As a microplate assay, the authors also propose the application of the test in early screening for genotoxic liability, as the qualities of high throughput and low compound use.</p>
<p>Hastwell PW, Webster TW, Tate M, Billinton N, Lynch AM, Harvey JS, Rees RW & Walmsley RM (2009). Analysis of 75 marketed pharmaceuticals using the GADD45a-GFP 'GreenScreen HC' genotoxicity assay. <i>Mutagenesis</i> 24: 455-463.</p>	<p>In this study, a collection of 75 marketed pharmaceuticals were tested in the GADD45a-GFP (GreenScreen HC) reporter assay which, detects genotoxic damage in the human lymphoblastoid TK6 cell line, giving positive results for all classes of genotoxins, including mutagens, aneugens and clastogens.</p>
<p>Jagger C, Tate M, Cahill PA, Hughes C, Knight AW, Billinton N & Walmsley RM (2009). Assessment of the genotoxicity of S9-generated metabolites using the GreenScreen HC GADD45a-GFP assay. <i>Mutagenesis</i> 24: 35-50.</p>	<p>The paper describes the method development, the derivation of decision thresholds for the identification of genotoxins using the method, and presents data from a 56-compound validation study of the method. The results illustrate that the method permitted the detection of the majority of pro-genotoxins tested and, importantly, the high specificity of the GADD45a-GFP assay was maintained.</p>
<p>Birrell L, Cahill P, Hughes C, Tate M & Walmsley RM (2010). GADD45a-GFP GreenScreen HC assay results for the ECVAM recommended lists of genotoxic and non-genotoxic chemicals for assessment of new genotoxicity tests. <i>Mutation Research</i> 695: 87-95.</p>	<p>Three categories of test chemicals comprising a total of 62 compounds were chosen based on the ECVAM recommended list of chemicals and analysed using the GreenScreen HC assay. The outcome results were in line with those from other studies with the GreenScreen HC assay confirming its high specificity.</p>
<p>Billinton N, Bruce S, Hansen JR, Hastwell PW, Jagger C, McComb C, Klug ML, Pant K, Rabinowitz A, Rees R, Tate M,</p>	<p>A new protocol has recently been developed for the GreenScreen HC GADD45a-GFP genotoxicity reporter assay, enabling the incorporation of an S9 metabolic activation system into</p>

Vinggaard AM & Walmsley RM (2010). A pre-validation transferability study of the GreenScreen HC GADD45a-GFP assay with a metabolic activation system (S9). <i>Mutation Research</i> 700: 44-50.	the assay. The outcome of the study was a refined protocol that was found to be highly transferable and reproducible between laboratories.
Hughes C, Rabinowitz A, Tate M, Birrell L, Allsup J, Billinton N & Walmsley RM (2012). Development of a high-throughput Gaussia luciferase Reporter Assay for the Activation of the GADD45a Gene by Mutagens, Promutagens, Clastogens, and Aneugens. <i>Journal of Biomolecular Screening</i> 17:1302-1316.	This article describes the development and validation of an alternative assay ("BlueScreen HC"), in which expression is linked to Gaussia luciferase (GLuc) expression, yielding a luminescent reporter, the preferred optical output in high-throughput screening. The new GLuc assay was as effective as the GFP (GreenScreen) assay in producing positive results for all classes of genotoxic carcinogen and negative results for all nongenotoxins tested.
Westerink WM, Stevenson JC, Horbach GJ & Schoonen WG (2010). The development of RAD51C, Cystatin A, p53 and Nrf2 luciferase-reporter assays in metabolically competent HepG2 cells for the assessment of mechanism-based genotoxicity and of oxidative stress in the early research phase of drug development. <i>Mutation Research</i> 696: 21-40.	Four different mechanism-based high-throughput luciferase-reporter assays were developed in human HepG2 cells. The promoter regions of RAD51C and Cystatin A, as well as the responsive element of the p53 protein, were selected for the generation of the genotoxicity reporter assays. Moreover, a luciferase-based reporter assay was generated that measures the activation of the Nrf2 oxidative stress pathway. Validation with respect to the ECVAM compound list resulted in an overall sensitivity of the HepG2 genotoxicity reporter assays for genotoxicity of 85%. The use of these assays in combination with the previously validated Vitotox and RadarScreen assays has been shown to be valuable to reducing the attrition rate due to genotoxicity in the developmental phase of drug development.
Testing Strategies	
Jaworska J, Gabbert S & Aldenberg T (2010). Towards optimization of chemical testing under REACH: a Bayesian network approach to Integrated Testing Strategies. <i>Regulatory Toxicology & Pharmacology</i> 57:157-167.	This paper presents conceptual requirements for ITS development and optimization. First, ITS development should be based on probabilistic methods in order to quantify and update various uncertainties across testing stages. Second, reasoning should reflect a set of logic rules for consistently combining probabilities of related events. Third, inference should be hypothesis-driven and should reflect causal relationships in order to coherently guide decision-making across testing stages. To meet these requirements, an information-theoretic approach to ITS development, the "ITS inference framework", which can be made operational by using Bayesian networks.
Kirkland D, Reeve L, Gatehouse D & Vanparys P (2011). A core <i>in vitro</i> genotoxicity battery comprising the Ames test plus the <i>in vitro</i> micronucleus test is sufficient to detect rodent carcinogens and <i>in vivo</i> genotoxins. <i>Mutation Research</i> 721: 27-73.	The paper presents an analysis of an existing database of rodent carcinogens and a new database of <i>in vivo</i> genotoxins in terms of the <i>in vitro</i> genotoxicity tests needed to detect their <i>in vivo</i> activity. The outcome of the analysis shows that there is no convincing evidence that any genotoxic rodent carcinogens or <i>in vivo</i> genotoxins would remain undetected in an <i>in vitro</i> two test battery consisting of Ames + MNvit.

<p>Kirkland D & Speit G (2008). Evaluation of the ability of a battery of three <i>in vitro</i> genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing <i>in vivo</i>. Mutation Research 654: 114-132.</p>	<p>The paper focuses on the discussion regarding the most appropriate follow-up testing <i>in vivo</i> when positive results are obtained <i>in vitro</i> but the <i>in vivo</i> micronucleus (MN) test (traditionally the most widely-used test) is negative. With the increased evaluation and use of other <i>in vivo</i> assays, e.g. for transgenic mutations (TG) and DNA damage (Comet assay) the authors considered extremely important to investigate their usefulness. The work examined the published <i>in vivo</i> UDS, TG and Comet-assay results for 67 carcinogens that were negative or equivocal in the micronucleus test. The data obtained suggested that Comet and TG assays should play a more prominent role in regulatory testing strategies than the UDS test.</p>
<p>Parry JM, Parry E, Phrakonkham P & Corvi R (2010). Analysis of published data for top concentration considerations in mammalian cell genotoxicity testing. Mutagenesis 25(6): 531-8.</p>	<p>The authors established a database of 384 chemicals classified as rodent carcinogens and reported Ames test results and the test concentrations that produced positive results in the mouse lymphoma assay (MLA), <i>in vitro</i> chromosome aberration (CA) assay and <i>in vitro</i> micronucleus test. This, in order to assess the impact that the reduction of testing concentrations would have on the outcome of <i>in vitro</i> genotoxicity testing. 62.5% produced positive results in the MLA, of which 20.3% required testing between 1 and 10 mM. A total of 58.0% produced positive results in <i>in vitro</i> CA assays, of which 25.0% required testing between 1 and 10 mM. If the testing concentration limit for mammalian cell assays was reduced to 1 mM, 24 (6.25%) potential carcinogens would not be detected in any part of the standard <i>in vitro</i> genotoxicity test battery (Ames test, MLA and <i>in vitro</i> CA assay). Further evaluation and/or retest of these compounds suggested that the current 10 mM top concentration can be reduced without any loss of sensitivity in detecting rodent carcinogens.</p>
<p>Kirkland D & Fowler P (2010). Further analysis of Ames-negative rodent carcinogens that are only genotoxic in mammalian cells <i>in vitro</i> at concentrations exceeding 1 mM, including retesting of compounds of concern. Mutagenesis 25 (6): 539-53.</p>	<p>In this article, the authors further evaluate the analysis performed by Parry and co-workers (Parry et al., 2010), regarding data for top concentration considerations in mammalian cell genotoxicity testing. The analysis focused on the 24 carcinogens which showed negative results in the Ames test and were positive in mammalian cells tests, instead, at concentrations between 1-10 mM. A more detailed evaluation of the studies performed and protocols used, many of them very old, prompted the authors to conclude that the 10 mM upper limit in mammalian cell tests could be lowered without any loss of sensitivity in detecting genotoxic rodent carcinogens. A new limit of 1 mM or 500 µg/ml was proposed.</p>
<p>Fowler P, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhrer S & Carmichael P (2012a). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. Mutation Research 742 (1-2):11-25.</p>	<p>Fowler and co-workers have compared several rodent cell lines (V79, CHL, CHO) with p53-competent human peripheral blood lymphocytes (HuLy), TK6 lymphoblastoid cells, and the human liver cell line HepG2, since p53-deficiency in many of the rodent cell lines has been considered a key factor in the poor testing predictivity. The authors evaluated <i>in vitro</i> micronucleus (MN) induction following treatment with 19 compounds that were accepted as</p>

	<p>producing misleading or "false" positive results in <i>in vitro</i> mammalian cell assays. The rodent cell lines (V79, CHO and CHL) resulted consistently more susceptible to cytotoxicity and MN induction than p53-competent cells, and were therefore more susceptible to giving misleading positive results. The authors suggested that a reduction in the frequency of misleading positive results could be achieved by careful selection of the mammalian cell type for genotoxicity testing.</p>
<p>Fowler P, Smith R, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael P (2012b). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. <i>Mutation Research</i> 747(1):104-17.</p>	<p>In line with the outcomes of the analysis mentioned in the above paper (Fowler et al., 2012a); the authors investigated the impact of different toxicity measures, commonly used <i>in vitro</i> cytogenetic assays, on the occurrence of misleading positive results. The data showed that estimating toxicity by relative cell count (RCC) or replication index (RI) consistently underestimated the toxicity observed by other measures (Relative Population Doubling, RPD, or Relative Increase in Cell Count, RICC). RCC and RI were more likely to lead to selection of concentrations for micronucleus scoring that were highly cytotoxic and thus could potentially lead to artefacts of toxicity being scored (elevated levels of apoptosis and necrosis), generating misleading positive results. The authors suggested that a further reduction in the frequency of misleading positive results in <i>in vitro</i> cytogenetic assays could be achieved (with 19 chemicals clearly giving misleading results) by avoiding the use of toxicity measures that underestimate the level of toxicity induced.</p>
<p>Pfuhler S, Kirkland D, Kasper P, Hayashi M, Vanparys P, Carmichael P, Dertinger S, Eastmond D, Elhajouji A, Krul C, Rothfuss A, Schoening G, Smith A, Speit G, Thomas C, van Benthem J & Corvi R (2009). Reduction of use of animals in regulatory genotoxicity testing: Identification and implementation opportunities-Report from an ECVAM workshop. <i>Mutation Research</i> 680: 31-42.</p>	<p>Report on the European Centre for the Validation of Alternative Methods (ECVAM) Workshop held in Ranco, Italy from 24 to 25 June 2008. The objectives of the workshop were to discuss how to reduce the number of animals in standard genotoxicity tests, whether the application of smarter test strategies can lead to lower animal numbers, and how the possibilities for reduction can be promoted and implemented.</p>
<p>Bowen DE, Whitwell JW, Lillford L, Henderson D, Kidd D, Mc Garry S, Pearce G, Beevers C & Kirkland DJ (2011). Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test. <i>Mutation Research</i> 722 (1): 7-19.</p>	<p>Presented here are the results of an evaluation trial in which the bone-marrow and peripheral blood (via MicroFlow(®) flow cytometry) micronucleus tests (looking at potential chromosome breakage and whole chromosome loss) in developing erythrocytes or young reticulocytes were combined with the Comet assay (measuring DNA strand-breakage), in stomach, liver and blood lymphocytes. This within the scope of revised draft ICH guidelines (Draft ICH S2) to establish a combined multi-end point <i>in vivo</i> assay to alleviate the need for multiple <i>in vivo</i> assays, thereby reducing time, cost and use of animals.</p>

<p>Rothfuss A, O'Donovan M, De Boeck M, Brault D, Czich A, Custer L, Hamada S, Plappert-Helbig U, Hayashi M, Howe J, Kraynak A, van der Leede B, Nakajima M, Priestley C, Thybaud V, Saigo K, Sawant S, Shi J, Storer R, Struwe M, Vock E & Galloway S (2010). Collaborative study on 15 compounds in the rat liver Comet Assay integrated into 2- and 4-week repeat-dose studies. <i>Mutation Research</i> 702: 40-69.</p>	<p>A collaborative trial was conducted to evaluate the possibility of integrating the rat-liver Comet assay into repeat-dose toxicity studies. Fourteen laboratories from Europe, Japan and the USA tested fifteen chemicals. Laboratories provided liver Comet assay data obtained at the end of the long-term (2- or 4-week) studies together with an evaluation of liver histology. Most of the test compounds were also investigated in the liver Comet assay after short-term (1–3 daily) administration to compare the sensitivity of the two study designs. MN analyses were also conducted in bone marrow or peripheral blood for most of the compounds to determine whether the liver Comet assay could complement the MN assay for the detection of genotoxins after long-term treatment.</p>
<p>Rothfuss A, Honma M, Czich A, Aardema MJ, Burlinson B, Galloway, Hamada S, Kirkland D, Heflich RH, Howe J, Nakajima M, O'Donovan M, Plappert-Helbig U, Priestley C, Recio L, Schuler M, Uno Y & Martus HJ (2011). Improvement of <i>in vivo</i> genotoxicity assessment: Combination of acute tests and integration into standard toxicity testing. <i>Mutation Research</i> 723: 108–120.</p>	<p>An IWGT working group reviewed current requirements for <i>in vivo</i> genotoxicity testing of different chemical product classes and identified opportunities for combination and integration of genotoxicity endpoints for each class. They considered: (1) combination of acute micronucleus (MN) and Comet assays into a single study, (2) integration of MN assays into repeated-dose toxicity (RDT) studies, (3) integration of Comet assays into RDT studies, and (4) requirements for the top dose when integrating genotoxicity measurements into RDT studies. Whereas the working group overall considered the presented data were adequate to conclude that the combination/integration of MNT and Comet is scientifically justified for both acute and RDT studies, it was noted that most recommendations made were based on limited data and may need to be refined in the future.</p>

9. Carcinogenicity

Federica Madia, Raffaella Corvi & Andrew Worth

9.1 The traditional in vivo tests and their regulatory use

Substances are defined as carcinogenic if after inhalation, ingestion, dermal application or injection they induce (malignant) tumours, increase their incidence or malignancy, or shorten the time of tumour occurrence. It is generally accepted that carcinogenesis is a multi-hit/multi-step process from the transition of normal cells into cancer cells via a sequence of stages and complex biological interactions, strongly influenced by factors such as genetics, age, diet, environment, hormonal balance, etc.

The induction of cancer involves genetic alterations which can be induced directly or indirectly. Therefore, carcinogens have conventionally been divided into two categories according to their presumed mode of action: genotoxic carcinogens and non-genotoxic carcinogens. Genotoxic carcinogens have the ability to interact with DNA and/or the cellular apparatus (such as e.g. the spindle apparatus and topoisomerase enzymes) and thereby affect the integrity of the genome, whereas non-genotoxic carcinogens exert their carcinogenic effects through other mechanisms that do not involve direct alterations to the DNA (Loeb & Harris, 2008; Adler, et al., 2011; Benigni, 2012).

Based on the above principles and before embarking into carcinogenicity studies, the genotoxic hazard of all type of substances is usually assessed (see chapter on Genotoxicity and Mutagenicity). The two-year cancer bioassay in rodents is widely regarded as the 'Gold standard' to evaluate cancer hazard and potency, although it is generally known that this test has its limitations to predict human cancer risk (Gottmann et al., 2001; Knight et al., 2005, 2006; Alden et al., 2011). Carcinogenicity assessment requires the conduct of either a carcinogenicity bioassay (OECD TG 451, EC B.32) or a combined chronic toxicity/carcinogenicity test (OECD TG 453, EC B.33). These studies take several years to run (the in-life portion alone lasts 24 months) and are very costly (one million Euro/chemical, approximately). In addition, repeat-dose toxicity studies, as well as other in vitro and in vivo studies can be used to better understand the human relevance of findings in rodents (e.g. related to toxicokinetic properties or species-specific effects) or the mechanism and associated dose-response for a chemical (OECD,2012a; Luijten et al., 2012).

At present, carcinogenicity testing, under REACH “can be proposed by the Registrant or may be required by ECHA” for chemicals at high level tonnage of production ($\geq 1000\text{t/y}$), which show widespread dispersive use **or** frequent or long-term human exposure **and** are classified as mutagens category 3 (GHS category 2), **or** there is evidence of hyperplasia and/or preneoplastic lesions from repeated dose toxicity studies (EC no. 1907/2006). However, REACH also requires that carcinogenic substances at all tonnage levels be identified as substances of high concern, taking into account information from all available relevant sources (non-human and human non-testing and testing data), which can inform hazard identification, underlying modes of action, or carcinogenic potency. The Biocidal Products legislation (EU no. 528/2012) requires carcinogenicity studies, unless substances are classified as mutagens category 1A and 1B or based on other specific justifications. A combined carcinogenicity study and long-term repeat-dose toxicity in the rat, possibly through the oral route plus a carcinogenicity study in a second rodent species, the mouse, are usually performed. For Plant Protection Products (PPRs) (EU nos. 1107/2009b and 283/284-2013) or Veterinary Drugs (VICH, 2005), tests for carcinogenicity are always required in two different species. Pharmaceuticals are also tested for carcinogenic potential, although several restrictions regarding the class of compound, duration of treatment (at least 6 months), dose or

target population may apply (ICH S1A, 1995; S1B, 2009; S1C, 2008; S1, 2012). For cosmetic ingredients and products no in vivo test is permitted from March 2013 (EC no. 1223/2009; SCCS, 1501/2012).

9.2 Mechanistic understanding of the endpoint

For an adequate evaluation of the genotoxic potential of a compound, different endpoints (i.e. induction of gene mutation, structural and numerical chromosome alterations) have to be assessed, as each of these events has been implicated in carcinogenesis and heritable diseases. A number of well-established and regulatory accepted in vitro tests are in place however, the low level of specificity often leads to misleading results that need to be clarified by in vivo genotoxicity tests (Kirkland et al., 2005; Maurici et al., 2005; Kirkland et al., 2007; Fowler et al., 2012a,b; refer also to chapter on genotoxicity and mutagenicity).

Although it is generally accepted that major carcinogenic risk is related to genotoxic compounds, the potential risk related to non-genotoxic compounds must also be evaluated. Several non-genotoxic substances may also cause tumours by affecting regulation of gene expression and genomic stability through epigenetic mechanisms, such as alternative states of gene expression, complex hyper- or hypo-methylation of DNA, histone modifications, inducing changes in protein folding and nucleosomal remodelling, RNA interference (Lo and Sukumar, 2008; Sadikovic *et al.*, 2008). Also sustained cell proliferation and inhibition of gap-junction intercellular communication are possible non-genotoxic mechanisms. Although some of the major mechanisms behind non-genotoxic carcinogenicity are known, multiple mechanisms of action and the underlying cellular and molecular events are not completely understood.

Typically modes of action of non-genotoxic substances are related to carcinogenesis phases of promotion and progression but participation into initiation phase is also proposed (Hattis et al., 2009). The induction of tissue-specific toxicity (cytotoxicity) resulting in inflammation and regenerative hyperplasia, for instance, is one of the well-known mechanisms (Loeb & Harris, 2008). Induction of immunosuppression by chemicals is regarded as another significant non-genotoxic mechanism of cancer (Hernández et al., 2009). Oxidative stress in cells also results in non-genotoxic carcinogenesis as it is shown that cancer cells commonly have increased levels of reactive oxygen species (ROS) and that ROS can induce cell malignant transformation (López-Lázaro, 2010; Klaunig et al., 2010). Oxidative stress has been suggested to have some involvement in the mode of carcinogenic activity of peroxisome proliferators in rodent livers (Doull et al., 1999, Hernández et al., 2009). Many non-genotoxic carcinogens act via binding to receptors such as aryl hydrocarbon, nuclear and peroxisome-proliferator receptors (Hattis et al., 2009), thus affecting cell proliferation, apoptosis and intercellular communication. Other cellular targets are tyrosine kinase (TK), ion channel-coupled and G-protein-coupled receptors (Silva Lima & Van der Laan, 2000). Many endocrine active substances act by binding to receptors such as the oestrogen, progesterone, aryl hydrocarbon and thyroid hormone receptors, thereby inducing cell proliferation in their target organs (Silva Lima & Van der Laan, 2000; Zhang & Ho, 2011; Labrecque et al., 2012; ECHA, 2012).

Since the mechanisms by which non-genotoxic carcinogens cause tumours are often related to species-specific disturbances in the normal physiological control of cellular proliferation, survival and differentiation (Widschwendter & Jones, 2002, Baylin & Ohm, 2006, Esteller, 2007), the observed effects are not necessarily predictive for humans (Shanks et al., 2009). In this context, the International Programme on Chemical Safety (IPCS) of the World Health Organisation (WHO) has been leading a project to harmonize approaches to the assessment of risk from exposure to chemicals, including chemical carcinogenesis and a Unified Human

Cancer Relevance Framework (IPCS HRF) has been produced. This is based on the concept that for a given tumour produced by a given compound in the experimental animal it may be possible to postulate a specific mode of action (MOA), the relevance of which at animal level can be clearly defined with a description of the key events (measurable parameters) that are along the causal path. The establishment of the specific animal MOA together with the analysis of dose-response and temporal relationships, along with analyses of the strength, consistency and specificity of key events, tumour responses, and biological plausibility and coherence, dynamic and kinetic factors allows a thorough qualitative and quantitative comparison between experimental animals and humans. By using this analytical approach, the human relevance of MOAs in the animal model can be determined, resulting in a transparent evaluation of human cancer risk (IPCS, 2007).

9.3 Status of non-standard methods and Integrated Approaches

Research into alternative predictive models for carcinogenicity has focused on medium-term carcinogenic *in vivo* models (e.g. *in vivo* transgenic models), short-term *in vitro* and *in vivo* biological assays, as well as computational models.

At present, the challenge is to develop alternative test methods for both genotoxic and non-genotoxic carcinogenicity. The complexity of the carcinogenicity process renders it difficult especially when developing *in vitro* alternative test models able to mimic the whole process and when considering non-genotoxic carcinogens. This challenge is also heightened because of the number of potential target organs. Some key events of the carcinogenesis process can be investigated *in vitro*. However, it is expected that an integrated approach involving multiple *in vitro* models will be needed, but a better understanding of the entire process is also required before this will be possible (Benfenati et al., 2009; Adler et al., 2011; Paules et al., 2011). Scientific research is on-going to try to achieve this goal. In this regard, it is worth noting the effort shown by IARC to highlight the potential role of epigenetic phenomena in cancer etiology and to drive the attention of scientific and regulatory communities towards the incorporation of epigenetic mechanisms into carcinogen identification and evaluation (Herczeg et al., 2013).

9.3.1 In vitro methods

Mammalian cell culture systems may be used to detect phenotypic changes *in vitro* induced by chemical substances associated with malignant transformation *in vivo* (DiPaolo et al., 1969; Isfort et al., 1996; Matthews et al., 1993). The cell transformation assays (CTAs) have been shown to closely model some key stages of the *in vivo* carcinogenesis process (LeBoeuf et al., 1999) and to date they represent the only standardised *in vitro* tests that have the potential to detect both genotoxic and non-genotoxic carcinogens. Moreover, the CTA is faster and more cost efficient than the *in vivo* rodent carcinogenicity assay, providing a useful approach for screening of chemicals with respect to their carcinogenic potential. Although the use of CTA is not currently an explicit regulatory requirement, its application is mentioned in a number of guidance documents (SCCS, 1501/2012; Jacobson-Kram & Jacobs, 2005; ECHA, 2012; Pfuhler et al., 2010; EFSA, 2011; ICH 2009). In these guidance documents CTAs are considered to be used as screening methods or to provide additional useful information to more routinely employed tests for assessing carcinogenic potential or to (Vanparys et al., 2012).

A study with SHE (pH 6.7 and 7.0) and BALB/c 3T3 cells was coordinated by ECVAM to address issues of standardisation of the protocols, transferability and reproducibility (Vanparys et al., 2010; Corvi et al., 2012). The study outcome in combination with the extensive database on CTA performances previously summarized in the OECD Detailed

Review Paper no. 31 (OECD, 2007; EURL ECVAM, 2012), supported the development of an OECD TG on CTA in SHE cells (Draft OECD TG, 2013).

An EURL ECVAM recommendation has been published on another CTA based on the Bhas42 cell line, which offers a relatively higher throughput than other CTA variants (Poth et al., 2007; Sakai et al., 2011; EURL ECVAM, 2013). EURL ECVAM recommended that CTAs should not be used as stand-alone tests, but may provide useful information about possible genotoxic and non-genotoxic carcinogenicity potential for use in conjunction with other data to generate supporting information for hazard identification that can eventually contribute to the risk assessment.

The assay may thus be used for these purposes in the context of a weight of evidence approach. Depending on the extent of other information available from non-testing and testing approaches, it is conceivable that information on the presence or lack of transforming potential of chemicals generated with the CTA may be sufficient for decision-making to confirm or refute a suspicion of carcinogenicity, and may thus in specific cases allow waiving the use of the rodent bioassay. Based on available data related to sensitivity and specificity, both positive and negative results may be relevant in the context of carcinogenicity evaluation. For example, depending on other information available, in some cases a positive CTA result can add confidence that a substance has the potential to be carcinogenic, while in some other situations a negative result could hint towards an absence of effect. In other cases, the CTA may provide testing data that still require confirmatory testing (EURL ECVAM, 2012). Further work should focus on the development of conceptual frameworks to define optimal combinations of CTA data with complementary information sets to address carcinogenicity assessment specific to different sectorial needs.

CTAs are currently used for clarification of *in vitro* positive results from genotoxicity assays to be used in the weight of evidence assessment. Data generated by CTAs can be useful where genotoxicity data for a certain substance class have limited predictive capacity (e.g. aromatic amines), for investigation of compounds with structural alerts for carcinogenicity or to demonstrate differences or similarities across a chemical category. Also the tumour-promoting activity of chemicals can be investigated by the CTAs (Vanparys et al., 2012). Moreover, it has been suggested that the CTA assay can be used in defining possible thresholds (e.g. DNELs/DMELs) for transforming activity (Jacquet et al., 2012; Vanparys et al., 2012). Recently, some tiered testing approaches which include the use of the CTAs have been proposed (Benigni et al., 2013; Doktorova et al., 2012).

Different research methods, including *in vitro* methods using several cell types, are available to study a number of potential non-genotoxic mechanisms. For example, tests are available to measure oxidative stress (Klaunig et al., 2010) or to measure the inhibition of gap junction intercellular communication (GJIC) (Klaunig & Shi, 2009), both of which have been associated with a number of non-genotoxic carcinogens. However, these methods cannot currently be used to reliably predict carcinogenic potential rather they are focused on better understanding the mechanism for effects elicited by a compound.

9.3.2 Toxicogenomics

In vitro toxicogenomics approaches using *in vitro* systems have shown to reach 80-90% accuracy (Tsujiyama et al., 2006; Li et al., 2007; Le Fevre et al., 2007; Ellinger-Ziegelbauer et al., 2009; Guyton et al., 2009; Mathijs et al., 2010; Waters et al., 2010; Doktorova, 2012; Jennings et al., 2013) for predicting *in vivo* toxicity in rodents although the number of chemicals is still limited and may not represent the full spectrum of toxins. Further *in vitro* tests were evaluated in the frame of the EU-funded project carcinoGENOMICS which aimed

at developing toxicogenomics- *in vitro* tests to detect potential genotoxicants and carcinogens in liver, lung and kidney target organs using different types of cells (carcinoGENOMICS Project website: <http://www.carcinogenomics.eu/index.php?id=100>; Doktorova 2013). Interestingly, the application of toxicogenomics to *in vitro* CTA assay resulted in a gene signature predictive of carcinogenic potential (Rohrbeck et al., 2010; Ao et al., 2010), which can be regarded as providing more confidence on the regulatory relevance of CTA test results.

Most *in vivo* toxicogenomics studies on assessment of carcinogenicity, focus, but do not limit, on short-term rat studies and non-genotoxic hepatocarcinogenicity and were shown to predict 80-90% rodent carcinogenicity which is in the same range of accuracy as for *in vitro* toxicogenomics data (Ellinger-Ziegelbauer et al., 2008; Nie et al., 2006; Fielden et al., 2007; Stemmer et al., 2007; Nioi et al., 2008; Uehara et al., 2008; Jonker et al., 2009; Vinken et al., 2012). They result instrumental for screening purpose or for better elucidation of modes of action (MOA), in particular in areas where the rodent bioassay is always required. The *in vivo* toxicogenomics assays can be helpful for hazard assessment and by that they may lead to a substantial reduction in the number of bioassays and the number of animals in the remaining *in vivo* tests. Indeed, the number of animals required for toxicogenomics-based assays is at least 10-fold-smaller and the exposure periods are 26-fold-shorter (max 4 weeks instead of 2 years) than for the standard rodent bioassay (Guo et al., 2006; MAQC Consortium, 2006).

Although significant technical progress has been made in the last decade, the limitations of toxicogenomics-based tests are many, including: quantitative risk assessment is in its infancy; limited public accessibility of raw data; poor knowledge of the function of many genes in the prediction sets; lack of uniformity in study design (e.g. rodent species and strain, dose setting criteria, time points, repeats) and bio-informatics analyses; requirement of expensive equipment and specialized staff. Moreover, formal validation of these methods is lacking at present and their regulatory acceptability is being discussed.

9.3.3 *In vivo* testing

Short-term tests with transgenic mouse models (p53+/-, rasH2, Tg.AC, Xpa-/- and Xpa-/-p53+) are possible alternatives for the classical 2-year cancer bioassay (Tennant et al., 1995 and 1999; Ashby, 2001). The rationale for using transgenic mice in regulatory carcinogenicity testing is that transgenic mouse models may be more sensitive predictors of carcinogenic risk to humans. Indeed these transgenic mouse models have a reduced tumour latency period (6-9 months) to chemically-induced tumours (Marx, 2003). Although not a complete replacement of the rodent 2-year cancer bioassay, transgenic mouse models can be considered as refinement methods and may result in a significant reduction in the use of experimental animals (20-25 animals/sex/treatment group) (ILSI/HESI ACT 2001; Eastin, 1998; Bucher, 1998; Pritchard, 2003; de Vries et al., 2004). While these animal model systems have been shown to be promising for the detection of carcinogens, their sensitivity and specificity remains to be determined. Due to a relative lack of validation, these assays are generally not yet accepted as a full replacement for the rodent 2-year bioassay, but data from these models may be used in a weight-of-evidence analysis of chemical carcinogenicity. It should however be noted that the evaluation study carried out by ILSI/HESI (ILSI/HESI Act, 2001) and a consortium of academic, governmental and industrial laboratories, led to the initial acceptance by pharmaceutical regulatory agencies of three primary models: p53+/-, Tg.AC and rasH2 model. The use of such models has been endorsed for the assessment of pharmaceuticals in lieu of a second species full carcinogenicity bioassay in mice (ICH S1B, 2009).

9.3.4 QSAR models

To date, hundreds of qualitative structure-activity relationship (SAR) and quantitative structure-activity relationship (QSAR) model have been published in the literature for predicting genotoxicity and carcinogenicity (Benfenati et al. 2009; Serafimova et al. 2010). These methods have proven successful at predicting genotoxic potential, and are arguably as reliable and informative as the gene mutation test in bacteria. On the other hand, QSARs capable of detecting non-genotoxic carcinogenic potential are still at an early stage of development and only a few are available (Benigni & Bossa, 2008; Toropov et al., 2009). A number of structural alerts and characteristics of several types of non-genotoxic carcinogens have been summarised (Woo et al., 1995; Woo & Lai, 2003; Ferrari & Gini, 2010).

Software tools in the public domain include CAESAR, Toxtree, OncoLogic and LAZAR. The models included in Toxtree (Benigni et al., 2008) and the OECD Toolbox are rule-based. Commercial models include MultiCase, TOPKAT, HazardExpert, Derek and ToxSuite. Derek and HazardExpert can be used in combination with programs Meteor and MetabolExpert to predict the genotoxicity and carcinogenicity potential of metabolites as well as parent compounds.

QSAR models for carcinogenic potency prediction are less accurate. This is expected given the complexity of the carcinogenicity endpoint, and the fact that models do not explicitly include ADME (absorption, distribution, metabolism excretion) properties, which could be critical steps in the carcinogenic process. While QSARs cannot be used as stand-alone alternatives to traditional tests, the combined use of several models can provide useful support to the overall evaluation of carcinogenicity (Cronin et al., 2003; Eriksson et al., 2003; Contrera et al., 2007; Jaworska and Nikolova-Jeliazkova, 2007; Benigni and Bossa, 2008; Fjodorova et al., 2010; Wu et al., 2010; Benigni, 2012; ECHA, 2010a,b; OECD, 2010d).

In addition, the use of categories, read-across and/or TTC approaches can be useful in filling data gaps and adding confidence to the predictions generated by QSARs (Van Leeuwen et al., 2009). Furthermore, additional confidence can be provided by the combination of multiple QSARs and in vitro tests (e.g. Peer Consultation on Health Canada Draft Weight of Evidence Framework for Genotoxic Carcinogenicity 2005).

9.3.5 Application of the Threshold of Toxicological Concern approach to cancer endpoints

The Threshold of Toxicological Concern (TTC) is a scientifically-based risk assessment approach to establish acceptable exposure limits when sufficient chemical-specific toxicological information is lacking (US FDA, 1995). To apply the TTC approach, a reliable measure or estimate of the exposure level is required. TTC is a useful approach for prioritising the further assessment of trace contaminants and chemicals associated with very low consumer exposures (see chapter on repeated dose toxicity). It is used, or has been proposed for use, in various sectors, including pharmaceuticals, chemicals in food, and chemicals in consumer products (Munro et al., 1996; Blackburn et al., 2005; Kroes et al., 2007; Safford RJ, 2008; Carthew et al., 2009; Felter et al., 2009; JEFCA, 1996-1997; EFSA, 2004; EMEA, 2006 and 2008; SCCP, 2008). Bercu et al. (2010) have suggested the use of TTC in combination with QSAR tools to establish safe levels for genotoxic impurities (GTIs) in drug substances.

9.3.6 Extrapolating from other animal studies

The role of genotoxicity testing can be both qualitative (hazard assessment) and quantitative (risk assessment). A linear relationship has been reported between the lowest effective dose

(LED) for in vivo genotoxicity and the carcinogen dose descriptor T25 (the chronic daily dose, which gives 25% of the animals tumours at a specific tissue site) (Sanner and Dybing, 2005). Indeed, the LED was suggested after further evaluation, as useful in a semi-quantitative method for risk assessment of mutagens without a long-term carcinogenicity study. The above results have been further supported by investigations on the applicability of in vivo and more recently in vitro genotoxicity tests to estimate cancer potency using the benchmark dose approach (Hernandez et al., 2011; Hernandez et al., 2013).

In light of the 3Rs Principle, the human pharmaceutical industry is currently exploring the possibility to predict carcinogenicity on the basis of available data (“Putting Animal Welfare Principles and 3Rs into Action” - European Pharmaceutical Industry 2011 Report), without necessarily requiring a 2-year bioassay. Rat chronic toxicology studies are good predictors of negative outcome in two-year rat carcinogenicity studies. The results derived from 182 compounds across 242 rat chronic toxicology studies and 182 two-year rat carcinogenicity studies conducted by 13 pharmaceutical companies over several years were reviewed (Sistare et al, 2011). The predictivity on an organ-by-organ basis is poor, but the overall negative predictivity is very good on a whole animal basis. Therefore, in the absence of chronic toxicity, preneoplasia, genotoxicity or hormonal perturbation signals, there is no added value from conducting a 2-year bioassay study. The potential preneoplastic histopathologic effects seen in chronic (six months) rat toxicology studies may be predictive of tumour outcome in 2-year bioassays. This is an interesting approach applicable to the pharmaceuticals sector however it is not directly relevant for REACH, since it would not require carcinogenicity testing under the above conditions. It is nevertheless worth investigating whether information related to other systemic toxicity endpoints can be extrapolated to provide information on carcinogenic potential, and potentially waive cancer bioassays. Similarly, for cosmetic ingredients in the past safety assessment for non-genotoxic chemicals has been based on identification of a NO(A)EL from repeat-dose (mainly 90 day) toxicity studies which along with appropriate conservative safety factors, has been used to perform a risk assessment for these chemicals, including the risk for carcinogenicity.

9.4 Conclusions

The major challenge for non-standard methods in carcinogenicity assessment is the identification and characterisation of non-genotoxic carcinogens. The proportion of non-genotoxic versus genotoxic carcinogens in the environment is likely to increase in the future, since the scientific knowledge on DNA reactivity allows industrial chemists to design compounds without overly reactive moieties. In contrast to genotoxic carcinogens, for those chemicals shown to lack genotoxicity potential, it is generally assumed that there is a threshold and that the carcinogenic risk can be assessed based on data from repeated dose toxicity studies. Prior to the formation of tumours (generally seen only after long-term exposures), non-genotoxic carcinogens cause changes in normal physiological function and these adverse effects can be determined in a repeat-dose toxicity study. For those chemicals the protection of human health against the pre-cancerous lesion will also protect against cancer which is a secondary effect. It is noted though, that some of these non-genotoxic carcinogens, when not classified for any other property and not identified as such in repeated dose toxicity studies, will go unidentified. While hyperplasia and pre-neoplastic lesions can be observed in a repeated dose toxicity study (e.g. in a sub-chronic study), other effects such as altered gene expression, DNA-methylation or inhibition of gap-functional communication are not monitored, as such, in repeated dose toxicity studies. In order to fully exploit the in vivo experiment, the integration of such additional multiparametric analyses into standard in vivo assays should be considered.

Efforts are needed to develop a strategy for evaluating hazard by combining information across different systemic toxicity endpoints, rather than considering them individually. This integrated approach is expected to result in a set of options for waiving redundant testing, which are motivated by a mechanistic understanding of the toxicological effects and their inter-relationships.

To avoid animal-specific and biased results, an *in vitro* testing battery based on human cell or tissue models with relevant biomarkers is seen as the most optimal way to replace animal tests in non-genotoxic carcinogenic assessment. It is expected that there will be significant synergies between work to develop replacement tests for repeat-dose toxicity studies with tests to predict non-genotoxic carcinogens and quantitative thresholds of response. Despite the fact that some mechanisms behind non-genotoxic carcinogenicity are known, multiple unknown mechanisms of action and the insufficient knowledge of the cellular and molecular events have not yet allowed for the imminent implementation of a battery of *in vitro* tests that could predict and/or explain their carcinogenic potential in humans. Moreover, those *in vitro* tests that could have a role in a testing strategy, which covers non-genotoxic endpoints, have not been sufficiently standardised and validated.

Although some short-term *in vitro* tests (e.g. CTA, toxicogenomics-based tests) are available beyond the standard *in vitro* genotoxicity tests to support conclusions on cancer hazard identification, the *in vitro* short-term tests are not sufficient to fully replace the animal tests needed to perform risk assessment for carcinogenicity. However, for some chemical classes the available non-animal methods might be sufficient to rule out carcinogenic potential in a weight of evidence approach. On the other hand, for clear genotoxic chemicals, a possibility is to rely on clear positive genotoxic results and label the chemical as a possible carcinogen, as the default presumption would be that a genotoxic mechanism for carcinogenicity is likely.

Although there is still a considerable need to develop reliable QSARs (and other theoretical models) for carcinogenicity prediction, existing models could be useful, but not as stand-alone approaches. Instead, QSARs can be applied, for example in the context of the TTC approach, to prioritise further testing, and they can be used for hazard identification in the context of weight-of-evidence approaches, along with *in vitro* data, grouping and read-across.

9.5 References

- Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleijnans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485
- Alden CL, Lynn A, Bourdeau A, Morton D, Sistare FD, Kadambi VJ & Silverman L (2011). A critical review of the effectiveness of rodent pharmaceutical carcinogenesis testing in predicting for human risk. *Veterinary Pathology* 48 (3): 772-84.
- Ao L, Liu JY, Liu WB, Gao LH, Hu R, Fang ZJ, Zhen ZX, Huang MH, Yang MS & Cao J (2010). Comparison of gene expression profiles in BALB/c 3T3 transformed foci exposed to tumor promoting agents. *Toxicology in Vitro* 24: 430-438.
- Ashby J (2001). Expectations for transgenic rodent cancer bioassay models. *Toxicologic Pathology* 29 Suppl, 177-182.
- Baylin SB & Ohm JE (2006). Epigenetic gene silencing in cancer – a mechanism for early oncogenic pathway addition? *Nature Reviews Cancer* 6: 107-116.
- Benfenati E, Benigni R, DeMarini DM, Helma C, Kirkland D, Martin TM, Mazzatorta P, Ouédraogo-Arras G, Richard AM, Schilter B, Schoonen WGEJ, Snyder RD & Yang C (2009). Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives. *Journal of Environmental Science and Health Part C Environmental Carcinogenesis & Ecotoxicology Reviews* 27 (2): 57-90.
- Benigni R. (2012). Alternatives to the carcinogenicity bioassay for toxicity prediction: are we there yet? *Expert Opinion Drug Metabolism & Toxicology* 8 (4): 407-17.
- Benigni R & Bossa C (2008). Predictivity of QSAR. *Journal of Chemical Information & Modeling* 48: 971-980.
- Benigni R, Bossa C, Jeliaskova N, Netzeva N & Worth A (2008). The Benigni/Bossa rulebase for mutagenicity and carcinogenicity – a module of Toxtree. JRC Report EUR 23241 EN. <http://publications.jrc.ec.europa.eu/repository>
- Benigni R, Bossa C & Tcheremenskaia O (2013). In vitro cell transformation assays for an integrated, alternative assessment of carcinogenicity: a data-based analysis. *Mutagenesis* 28 (1): 107-16.
- Bercu JP, Morton SM, Deahl JT, Gombar VK, Callis CM & van Lier RB (2010). In silico approaches to predicting cancer potency for risk assessment of genotoxic impurities in drug substances. *Regulatory Toxicology & Pharmacology* 57: 300-306.
- Blackburn K, Stickney JA, Carlson-Lynch HL, McGinnis PA, Chappell L & Felter SP (2005). Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regulatory Toxicology & Pharmacology* 43: 249-259.
- Bucher JR (1998). Update on National Toxicology Program (NTP) assays with genetically altered or “transgenic” mice. *Environmental Health Perspectives* 106: 619-621.
- Cronin MT, Jaworska JS, Walker JD, Comber MH, Watts CD & Worth AP. (2003). Use of QSARs in

- International Decision-Making Frameworks to Predict Health Effects of Chemical Substances
Environmental Health Perspectives 111 (10): 1391-1401.
- Contrera JF, Kruhlak NL, Matthews EJ & Benz RD (2007). Comparison of MC4PC and MDL-QSAR rodent carcinogenicity predictions and the enhancement of predictive performance by combining QSAR models. *Regulatory Toxicology & Pharmacology* 49:172-182.
- Carthew P, Clapp C, & Gutsell S (2009). Exposure based waiving: the application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. *Food Chemical Toxicology* 47: 1287-1295.
- Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S, Schechtman L & Vanparys P. (2012). ECVAM prevalidation study on in vitro cell transformation assays: general outline and conclusions of the study. *Mutation Research* 744 (1): 12-9.
- DiPaolo JA, Donovan P & Nelson R (1969). Quantitative studies of in vitro transformation by chemical carcinogens. *Journal of the National Cancer Institute* 42: 867-874.
- Doull J, Cattley R, Elcombe C, Lake BG, Swenberg J, Wilkinson C, Williams G & van Gemert M (1999). A cancer risk assessment of di(2-ethylhexyl) phthalate: application of the new U.S. EPA risk assessment guidelines. *Regulatory Toxicology & Pharmacology* 29 (3): 327-357.
- De Vries A, van Steeg H & Opperhuizen A (2004). Transgenic mice as alternatives in carcinogenicity testing: current status. RIVM report 340700001.
- Doktorova TY, Pauwels M, Vinken M, Vanhaecke T & Rogiers V (2012) Opportunities for an alternative integrating testing strategy for carcinogen hazard assessment? *Critical Reviews in Toxicology* 42 (2): 91-106.
- Doktorova TY, Yildirimman R, Vinken M, Vilardell M, Vanhaecke T, Gmuender H, Bort R, Brolen G, Holmgren G, Li R, Chesne C, van Delft J, Kleinjans J, Castell J, Bjorquist P, Herwig R & Rogiers V (2013). Transcriptomic responses generated by hepatocarcinogens in a battery of liver-based in vitro models. *Carcinogenesis* 34: 1393-1402..
- EC (2006). Regulation No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. . OJ L 396 of 30122006: 1-849.
- EC (2009). Regulation No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products (Text with EEA relevance). OJ L 342 of 22122009.
- EC (2009b). Regulation No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309 of 24112009: 1-50.
- ECHA (2010a). Practical Guide on how to report (Q)SARs.
http://echa.europa.eu/doc/publications/practical_guides/pg_report_qsars.pdf
- ECHA (2010b). Practical Guide on how to report Categories and Read-Across.
http://echa.europa.eu/doc/publications/practical_guides/pg_report_readacross_categ.pdf
- ECHA (2012). Guidance, for the implementation of REACH. Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific Guidance Version 2.0 November 2012, R.7.7. Mutagenicity and Carcinogenicity.

EURL ECVAM (2012). Recommendation concerning the cell transformation assays using Syrian hamster embryo cells (SHE) and the BALB/c 3T3 mouse fibroblast cell line for in vitro carcinogenicity testing. Annex I: ESAC opinion on the ESAC peer review of an ECVAM-coordinated prevalidation study concerning three protocols of the cell transformation assay (CTA) for in vitro carcinogenicity testing. http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing

EURL ECVAM (2013) Recommendation on the Cell Transformation Assay based on the Bhas 42 cell line
http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-cell-transformation-assay-based-on-the-bhas-42-cell-line

EFSA (2004). European Food Safety Authority. Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food. Opinion on Flavouring Group FGE.03 Acetals of branched- and straight-chain aliphatic saturated primary alcohols and branched- and straight-chain saturated aldehydes, and an orthoester of formic acid, from chemical groups 1 and 2. Opinion expressed on 7 October 2004. EFSA J. 107, 1-59.

EFSA (2011). Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011; 9 (9): 2379.

Ellinger-Ziegelbauer H, Aubrecht J, Kleinjans JC & Ahr HJ (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. Toxicology Letters 186: 36-44.

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A & Ahr HJ (2008). Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term *in vivo* studies. Mutation Research 637: 23-39.

EMA (2006). European Medicines Evaluation Agency. Committee for Medicinal Products for Human Use (CHMP). Guideline on the limits of genotoxic impurities. CPMP/SWP/5199/02. London, 28 June 2006. EMA/CHMP/QWP/ 251344/2006.

EMA (2008). Committee for Medicinal Products for Human Use (CHMP). Questions & Answers on the CHMP Guideline on the limits of Genotoxic Impurities. Rev 1. London, 26 June 2008. Doc.Ref. EMA/CHMP/SWP/431994/2007

Esteller M (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nature Reviews Genetics 8: 286-298.

EU (2003). Directive 2003/15/EC of the European Parliament and the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Member States relating to cosmetic products. Official Journal of the European Union L66, 26-35.

EU (2012). Regulation No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products (Text with EEA relevance). OJ L 167 of 22052012: 1-123.

EU (2013). Commission Regulation No 283/2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93 of 01032013: 1-84.

EU (2013b). Commission Regulation No 284/2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market (Text with EEA relevance). OJ L 93 of 01032013: 85-152.

Eastin WC, Haseman JK, Mahler JF & Bucher JR (1998). The National Toxicology program evaluation of genetically altered mice as predictive models for identifying carcinogens. *Toxicologic Pathology* 26: 461-473.

Felter SP, Lane RW, Latulippe ME, Llewellyn GC, Olin SS, Scimeca JA & Trautman TD (2009). Refining the threshold of toxicological concern (TTC) for risk prioritization of trace chemicals in food. *Food and Chemical Toxicology* 47: 2236-2245.

Ferrari T & Gini G (2010). An open source multistep model to predict mutagenicity from statistical analysis and relevant structural alerts. *Chemistry Central Journal* 4 (Suppl 1), S2

Fielden MR, Brennan R & Gollub J (2007). A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicological Sciences* 99: 90-10.

Fjodorova N, Vracko M, Novic M, Roncaglioni A & Benfenati E (2010). New public QSAR model for carcinogenicity. *Chemistry Central Journal* 4 Suppl 1:S3

Fowler P, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael P (2012). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. *Mutation Research* 742 (1-2): 11-25.

Fowler P, Smith R, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael P (2012). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. *Mutation Research* 747 (1): 104-17.

Gottmann E, Kramer S, Pfahringer B & Helma C (2001). Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. *Environmental Health Perspectives* 109 (5): 509-14.

Guo L, Lobenhofer EK, Wang C, Shippey R, Harris SC, Zhang L, Mei N, Chen T, Herman D, Goodsaid FM, Hurban P, Philips KL, Qu J, Deng X, Sun YA, Tong W, Dragon WT & Shi L (2006). Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nature Biotechnology* 24: 1162-1169.

Guyton KZ, Kyle AD, Aubrecht J, Cogliano VJ, Eastmond DA, Jackson M, Keshava N, Sandy MS, Sonawane B, Zhang L, Waters MD & Smith MT (2009). Improving prediction of chemical carcinogenicity by considering multiple mechanisms and applying toxicogenomic approaches. *Mutation Research* 681: 230-240.

Hattis D, Chu M, Rahmioglu N, Goble R, Verma, P, Hartman K & Kozlak M (2009). A preliminary operational classification system for nonmutagenic modes of action for carcinogenesis. *Critical Reviews in Toxicology* 39 (2): 97-138.

Herceg Z, Lambert M-P, van Veldhoven K, Demetriou C, Vineis P, Smith MT, Straif K & Wild CP (2013). Towards incorporating epigenetic mechanisms into carcinogen identification and evaluation. *Carcinogenesis* doi:10.1093/carcin/btg212. 1-13.

Hernández LG, van Steeg H, Luijten M & van Benthem J (2009). Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutation Research* 682: 94-109.

Hernández LG, Slob W, van Steeg H & van Benthem J (2011). Can carcinogenic potency be predicted from *in vivo* genotoxicity data?: a meta-analysis of historical data. *Environmental & Molecular Mutagenesis* 52 (7): 518-28.

Hernández LG, van Benthem J & Johnson GE (2013). A mode-of-action approach for the identification of genotoxic carcinogens. *PLoS One* May 13; 8 (5): e64532.

- ICH (1995) S1A ICH harmonized Tripartite Guideline, Guideline on the Need for Carcinogenicity Studies of Pharmaceuticals, S1A, Current Step version 4, November 1995.
- ICH (2009) S1B Guideline on carcinogenicity testing of pharmaceuticals.
http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002735.pdf.
- ICH (2008) S1C(R2) Dose selection for carcinogenicity Studies of Pharmaceuticals, S1C (R2), Current Step version 4, 11 March 2008.
- ICH (2009) M3(R2) Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. June 2009.
- ICH Concept Paper (2012) S1: Rodent Carcinogenicity Studies for Human Pharmaceuticals, Dated and endorsed by the Steering Committee on 14 November 2012
- ICH (2012) S1: Regulatory notice on changes to core guideline on rodent carcinogenicity testing of pharmaceuticals. EMA/CHMP/ ICH/752486/2012 December 2012.
- ILSI HESI ACT (2001). ILSI HESI Alternatives to carcinogenicity testing project. Toxicologic Pathology 29 supplement: 1-351.
- IPCS (2007), Harmonization Project Document No. 4 Part 1: IPCS Framework for analysing the relevance of a cancer mode of action for humans and case-studies. IPCS, WHO
http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf
- Isfort RJ, Kerckaert GA & LeBoeuf RA (1996). Comparison of the standard and reduced pH Syrian hamster embryo (SHE) cell *in vitro* transformation assays in predicting the carcinogenic potential of chemicals. Mutation Research 356: 11-63.
- Jacobson-Kram D & Jacobs A (2005). Use of genotoxicity data to support clinical trials or positive genotox findings on a candidate pharmaceutical or impurity ... now what? International Journal of Toxicology. 24 (3): 129-34.
- Jacquet N, Maire MA, Landkocz Y, & Vasseur P (2012). Carcinogenic potency of perfluorooctane sulfonate (PFOS) on Syrian hamster embryo (SHE) cells. Archives of Toxicology 86 (2): 305-14.
- Jaworska J & Nikolova-Jeliazkova N (2007). How can structural similarity analysis help in category formation? SAR and QSAR in Environmental Research 18: 195-207.
- JECFA (1996). A procedure for the safety evaluation of flavouring agents. WHO Food Additive Series 35. Annex 5. World Health Organization, Geneva
- JECFA (1997). Safety evaluation of certain food additives and contaminants. Forty-ninth Report of the Joint FAO/WHO Expert Committee on Food Additives. WHO Food Additive Series 40. World Health Organization, Geneva
- Jennings P, Limonciel A, Felice L & Leonard MO (2013). An overview of transcriptional regulation in response to toxicological insult. Archives of Toxicology 87: 49–72.
- Jonker MJ, Bruning O, van Itersen M, Schaap MM, van der Hoeven TV, Vrieling H, Beems RB, de Vries A, van Steeg H, Breit TM & Luijten M (2009). Finding transcriptomics biomarkers for *in vivo* identification of (non-)genotoxic carcinogens using wild-type and Xpa/p53 mutant mouse models. Carcinogenesis 30: 1805-1812.
- Kirkland D, Aardema M, Henderson L & Müller L (2005). Evaluation of the ability of a battery of

- three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutation Research* 584: 1-256.
- Kirkland D, Pfulher S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier J-R, Müller L, Nohynek G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P & White P (2007). How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. *Mutation Research* 628: 31-55.
- Klaunig JE & Shi Y (2009). Assessment of gap junctional intercellular communication. *Current Protocols in Toxicology*. Chapter 2:Unit 2.17.
- Klaunig JE, Kamendulis LM & Hocevar BA (2010). Oxidative stress and oxidative damage in carcinogenesis. *Toxicologic Pathology* 38 (1): 96-109.
- Knight A, Jonker MJ & Bruning J (2005). Which drugs cause cancer? *British Medical Journal USA* 5: 477-478.
- Knight A, Bailey J & Balcombe J (2006). Animal carcinogenicity studies: Implications for the REACH system. *Alternatives to Lab Animals* 34, Suppl. 1: 139-147.
- Kroes R, Renwick AG, Feron V, Galli CL, Gibney M, Greim H, Guy RH, Lhuguenot JC & Van de Sandt JJM (2007). Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food & Chemical Toxicology* 45: 2533-2562.
- Labrecque MP, Takhar MK, Hollingshead BD, Prefontaine GG, Perdew GH & Beischlag TV (2012). Distinct roles for aryl hydrocarbon receptor nuclear translocator and ah receptor in estrogen-mediated signalling in human cancer cell lines. *PLoS One* 7 (1): e29545.
- LeBoeuf RA, Kerckaert KA, Aardema MJ & Isfort RJ (1999). Use of Syrian hamster embryo and BALB/c 3T3 cell transformation for assessing the carcinogenic potential of chemicals. *IARC Science Publications* 146: 409-25.
- Le Fevre AC, Boitier E, Marchandeu JP, Sarasin A & Thybaud V (2007). Characterization of DNA reactive and non-DNA reactive anticancer drugs by gene expression profiling. *Mutation Research* 619: 16-29.
- Li HH, Aubrecht J & Fornace AJ Jr (2007). Toxicogenomics: overview and potential applications for the study of non-covalent DNA interacting chemicals. *Mutation Research* 623: 98-108.
- Lo PK & Sukumar S (2008). Epigenomics and breast cancer. *Pharmacogenomics* 9 (12): 1879-902.
- Loeb LA & Harris CC. (2008). Advances in chemical carcinogenesis: a historical review and prospective. *Cancer Research* 68 (17): 6863-72.
- López-Lázaro M (2010). A new view of carcinogenesis and an alternative approach to cancer therapy. *Molecular Medicine* 16 (3-4): 144-153.
- Luijten, M, Muller JJA, Hernández LG, van der Ven LTM & van Benthem J (2012). Prediction of carcinogenic potential of substances using repeated dose toxicity data. RIVM Report 340700006, pp 1-18.
- Mathijs K, Brauers KJ, Jennen DGJ, Lizarraga D, Kleinjans JCS & van Delft JHM (2010). Gene expression profiling in primary mouse hepatocytes discriminates true from false-positive genotoxic compounds. *Mutagenesis* 25: 561-568.

Matthews EJ, Spalding JW & Tennant RW (1993). Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays. *Environmental Health Perspectives* 101 Suppl 2: 347-482.

MAQC Consortium (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnology* 24: 1151-1161.

Maurici D, Aardema M, Corvi R, Kleber M, Krul C, Laurent C., Loprieno N, Pasanen M, Pfuhler S, Phillips B, Sabbioni E, Sanner T & Vanparys P (2005). Genotoxicity and mutagenicity. *Alternatives to Lab Animals* 33 Suppl 1: 117-130.

Marx J (2003). Building Better Mouse Models for Studying Cancer. *Science* 299: 1972-1975.

Munro IC, Ford RA, Kennepohl E & Sprenger JG (1996). Correlation of structural class with non-observed effect levels: a proposal for establishing a threshold of concern. *Food & Chemical Toxicology* 34: 829-867.

Nie AY, McMillian M, Parker JB, Leone A, Bryant S, Yieh L, Bittner A, Nelson J, Carmen A, Wan J & Lord PG (2006). Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Molecular Carcinogenesis* 45: 914-933.

Nioi P, Pardo ID, Sherratt PJ & Snyder RD (2008). Prediction of non-genotoxic carcinogenesis in rats using changes in gene expression following acute dosing. *Chemico-Biological Interactions* 172: 206-215.

OECD (2007a) Detailed Review Paper no. 31. On Cell Transformation Assays for Detection of Chemical Carcinogens, ENV/JM/MONO (2007)18

OECD (2007b). Guidance Document no. 80. On Grouping of Chemicals, OECD Environment Health and Safety Publications, Series on Testing and Assessment. [http://apli1.oecd.org/olis/2007doc.nsf/linkto/env-jm-mono\(2007\)28](http://apli1.oecd.org/olis/2007doc.nsf/linkto/env-jm-mono(2007)28)

OECD (2012a) Guidance Document, no. 116. On the conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Support test Guidelines 451, 452 and 453 2ND Edition, ENV/JM/MONO (2011) 47; 13-Apr-2012

OECD (2013) DRAFT TG Publication: In vitro Carcinogenicity: Syrian Hamster Embryo (SHE) Cell Transformation Assay. October 2012.
<http://www.oecd.org/env/ehs/testing/Draft%2017%20October%202012.pdf>

Paules RS, Aubrecht J, Corvi R, Garthoff B & Kleijnans JC (2011). Moving Forward in Human Cancer Risk Assessment. *Environmental Health Perspectives* 119:439-443.

Pauwels M & Rogiers V (2009). The safety evaluation of cosmetic ingredients at the EU level and the use of 3R-alternatives: an update. Poster presentation in the "7th World Congress on Alternatives & Animal Use in the Life Sciences", Rome, Italy, ALTEX 26 (Special Issue): 193.

Pfuhler S, Kirst A, Aardema M, Banduhn N, Goebel C, Araki D, Costabel-Farkas M, Dufour E, Fautz R, Harvey J, Hewitt NJ, Hibatallah J, Carmichael P, Macfarlane M, Reisinger K, Rowland J, Schellauf F, Schepky A & Scheel J (2010). A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: Genotoxicity. A COLIPA analysis. *Regulatory Toxicology & Pharmacology* 57: 315-324.

Peer Consultation on Health Canada Draft Weight of Evidence Framework for Genotoxic Carcinogenicity (2005).
<http://www.tera.org/peer/comhaz/HC%20QSAR%20Peer%20Consultation%20Overview.pdf>

Poth A, Heppenheimer A & Bohnenberger S. (2007). Bhas42 cell transformation assay as a predictor

of carcinogenicity. ALTEX. 14: 519-521.

Pritchard JB, French JE, Davis BJ & Haseman JK (2003). The role of transgenic mouse models in carcinogen identification. Environmental Health Perspectives 111: 444-454.

Rohrbeck A, Salinas G, Maaser K, Linge J, Salovaara S, Corvi R & Borlak J (2010). Toxicogenomics applied to *in vitro* carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens. Toxicological Sciences 118: 31-41.

Sadikovic B, Al-Romaih K, Squire JA & Zielenska M (2008). Cause and consequences of genetic and epigenetic alterations in human cancer. Current Genomics 9: 394-408.

Safford RJ (2008). The dermal sensitization threshold - A TTC approach for allergic contact dermatitis. Regulatory Toxicology & Pharmacology 51: 195-200.

Sakai A, Sasaki K, Hayashi K, Muramatsu D, Arai S, Endou N, Kuroda S, Albrecht P, Bohnenberger S, Kunkelmann T, Asakura M, Hirose H, Ishii N, Mizuhashi F, Kasamoto S, Nagai M, Pant K, Bruce SW, Sly JE, Yamazaki S, Umeda M & Tanaka N (2011). An international validation study of a Bhas 42 cell transformation assay for the prediction of chemical carcinogenicity. Mutation Research 725: 57-77.

Sanner T & Dybing E (2005). Comparison of carcinogenic and *in vivo* genotoxic potency estimates. Basic Clinical Pharmacology & Toxicology 96: 131-139.

SCCS (2012) Notes of Guidance for the testing of cosmetics substances and their safety evaluation, 8th Revision, SCCS/1501/12.

SCCP (2008) Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products SCCP/1171/08 (SCCS/SCHER/SCENIHR adopted on June 8, 2012).

Serafimova R, Fuat Gatnik M & Worth A (2010). Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC report EUR 24427 EN.
<http://publications.jrc.ec.europa.eu/repository>

Serafimova R, Todorov M, Pavlov T, Kotov S, Jacob E, Aptula A & Mekenyan O (2007). Identification of the structural requirements for mutagenicity, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. Chemical Research in Toxicology 20: 662-676.

Shanks N, Greek R & Greek J (2009). Are animal models predictive for humans? Philosophy, Ethics & Humanities in Medicine 4: 1-20.

Silva Lima B & Van der Laan JW (2000). Mechanisms of nongenotoxic carcinogenesis and assessment of the human hazard. Regulatory Toxicology & Pharmacology 32 (2): 135-43.

Sistare FD, Morton D, Alden C, Christensen J, Keller D, Jonghe SD, et al. (2011). An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. Toxicologic Pathology 39 (4): 716-44.

Stemmer K, Ellinger Ziegelbauer H, Ahr HJ & Dietrich DR (2007). Carcinogen-specific gene expression profiles in short-term treated Eker and wild-type rats indicative of pathways involved in renal tumorigenesis. Cancer Research 67: 4052-4068.

Tennant RW, French JE & Spalding JW (1995). Identifying chemical carcinogens and assessing potential risk in short-term bioassays using transgenic mouse models. Environmental Health Perspectives 103: 942-950.

Tennant RW, Stasiewicz S, Mennear J, French JE & Spalding JW (1999). Genetically altered mouse models for identifying carcinogens. IARC Scientific Publications 146: 123-150.

Toropov AA, Toropova AP & Benfenati E (2009). Additive SMILES-based carcinogenicity models: probabilistic principles in the search for robust predictions. International Journal of Molecular Science 10: 3106-3127.

Tsujimura K, Asamoto M, Suzuki S, Hokaiwado N, Ogawa K & Shirai T (2006). Prediction of carcinogenic potential by a toxicogenomic approach using rat hepatoma cells. Cancer Science 97: 1002-1010.

Uehara T, Hirode M, Ono A, Kiyosawa N, Omura K, Shimizu T, Mizukawa Y, Miyagishima T, Nagao T & Urushidani T (2008). A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. Toxicology 250: 15-26.

U.S. FDA (1995) Department of Health and Human Services. Food Additives; Threshold of Regulation for Substances Used in Food-Contact Articles 21 CFR Parts 5, 25, 170, 171, and 174

Van Leeuwen K, Schultz TW, Henry T, Diderich B & Veith GD (2009). Using chemical categories to fill data gaps in hazard assessment. SAR QSAR Environmental Research 20: 207-220.

Vanparys Ph, Corvi R, Aardema M, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2010). ECVAM prevalidation of three cell transformation assays. ALTEX 27: 267-270.

Vanparys P, Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2012). Application of *in vitro* cell transformation assays in regulatory toxicology for pharmaceuticals, chemical, food products and cosmetics. Mutation Research 744 (1): 111-116.

VICH (2005) GL28, Safety Carcinogenicity. Studies to evaluate the safety of residues of veterinary drugs in human food: Carcinogenicity Testing. February 2005, for implementation at Step 7 – Final (Chapter 2.)

Vinken M, Vanhaecke T & Rogiers V (2012). Primary hepatocyte cultures as *in vitro* tools for toxicity testing: quo vadis? Toxicology *in Vitro* 26: 541-544.

Waters MD, Jackson M & Lea I (2010). Characterizing and Predicting Carcinogenicity and Mode of Action Using Conventional and Toxicogenomics Methods. Mutation Research 705: 184-200.

Widschwendter M & Jones PA (2002). DNA methylation and breast carcinogenesis. Oncogene 21: 5462-5482.

Woo YT & Lai DY (2003). Mechanisms of action of chemical carcinogens, and their role in Structure-Activity Relationships (SAR) analysis and risk assessment. In: Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens (R Benigni, ed). Boca Raton: CRC Press: 41-80.

Woo YT, Lai DY, Argus MF & Arcos JC (1995). Development of structure–activity relationship rules for predicting carcinogenic potential of chemicals. Toxicology Letters 79: 219-228.

Wu S, Blackburn K, Amburgey J, Jaworska J & Federle T (2010). A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regulatory Toxicology & Pharmacology 56: 67-81.

Zhang X & Ho SM (2011). Epigenetics meets endocrinology. Journal of Molecular Endocrinology 46 R11-R32.

Chapter 9. Table of References with Notes

Traditional <i>in vivo</i> tests	
<p>OECD (2009), Test No. 451: Carcinogenicity Studies, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071186-en</p>	<p>The objective of a long-term carcinogenicity study is to observe test animals for a major portion of their life span for the development of neoplastic lesions during or after exposure to various doses of a test substance by an appropriate route of administration.</p> <p>This Test Guideline is intended primarily for use with rats and mice, and for oral administration. Both sexes should be used. Each dose group and concurrent control group should contain at least 50 animals of each sex. At least three dose levels and a concurrent control should be used. Animals are dosed with the test substance daily (oral, dermal or inhalation administration) and the mode of exposure should be adjusted according to the toxicokinetic profile of the test substance. The duration of the study will normally be 24 months for rodents. For specific strains of mice, duration of 18 months may be more appropriate. Termination of the study should be considered when the number of survivors in the lower dose groups or the control group falls below 25 per cent. The results of these studies include: measurements (weighing, food consumption), and, at least, daily and detailed observations, as well as gross necropsy and histopathology.</p>
<p>OECD (2009), Test No. 452: Chronic Toxicity Studies, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071209-en</p>	<p>The objective of these chronic toxicity studies is to characterize the profile of a substance in a mammalian species (primarily rodents) following prolonged and repeated exposure.</p> <p>The Test Guideline focuses on rodents and oral administration. Both sexes should be used. For rodents, at least 20 animals per sex per group should normally be used at each dose level, while for non-rodents a minimum of 4 per sex per group is recommended. At least three dose levels should be used in addition to the concurrent control group. Frequency of exposure normally is daily, but may vary according to the route chosen (oral, dermal or inhalation) and should be adjusted according to the toxicokinetic profile of the test substance. The duration of the exposure period should be 12 months. The study report should include: measurements (weighing) and regular detailed observations (haematological examination, urinalysis, clinical chemistry), as well as necropsy procedures and histopathology.</p>
<p>OECD (2009), Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. doi: 10.1787/9789264071223-en</p>	<p>The objective of a combined chronic toxicity/carcinogenicity study is to identify carcinogenic and the majority of chronic effects, and to determine dose-response relationships following prolonged and repeated exposure.</p> <p>The rat is typically used for this study. For rodents, each dose group and concurrent control group intended for the carcinogenicity phase of the study should contain at least 50 animals of each sex, while for the chronic toxicity phase of the study should contain at least 10 animals of each sex. At least three dose levels should be used, in addition to the concurrent control group for both the</p>

	<p>chronic toxicity phase and the carcinogenicity phase of the study. The three main routes of administration are oral, dermal, and inhalation. The Test Guideline focuses on the oral route of administration.</p> <p>The period of dosing and duration of the study is normally 12 months for the chronic phase, and 24 months for the carcinogenicity phase. The study report should include: measurements (weighing) and regular detailed observations (haematological examination, urinalysis, clinical chemistry), as well as necropsy procedures and histopathology. All these observations permit the detection of neoplastic effects and a determination of carcinogenic potential as well as the general toxicity.</p>
New <i>in vivo</i> and <i>in vitro</i> tests – Draft OECD TGs	
<p>OECD (2013) <i>DRAFT</i> TG Publication: In vitro Carcinogenicity: Syrian Hamster Embryo (SHE) Cell Transformation Assay. October 2012 http://www.oecd.org/env/ehs/testing/Draft%2017%20October%202012.pdf</p>	<p><i>In vitro</i> cell transformation refers to the induction of phenotypic alterations in cultured cells that are characteristic of tumorigenic cells. This Test Guideline (TG) provides an <i>in vitro</i> procedure of the cell transformation assay in Syrian Hamster Embryonic stem cells (SHE), which may be used for hazard identification of chemical carcinogens.</p>
<p>OECD (2013) <i>DRAFT</i> TG Publication: <i>In vivo</i> mammalian Alkaline Comet Assay. http://www.oecd.org/env/ehs/testing/OECD_draft%20TG_Comet_December2013.pdf</p>	<p>Rodent <i>in vivo</i> Comet assays are especially relevant to assessing DNA-damaging potential of chemicals for regulatory purposes.</p>
<p>EURL ECVAM (2012) Recommendation concerning the cell transformation assays using Syrian hamster embryo cells (SHE) and the BALB/c 3T3 mouse fibroblast cell line for <i>in vitro</i> carcinogenicity testing. Annex I: ESAC opinion on the ESAC peer review of an ECVAM-coordinated prevalidation study concerning three protocols of the cell transformation assay (CTA) for <i>in vitro</i> carcinogenicity testing. http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing</p>	<p>EURL ECVAM recommendation report on the three CTA protocol variants: (a) The SHE CTA performed at pH 6.7 (SHE pH 6.7 CTA); (b) the SHE CTA performed at pH 7.0 (SHE pH 7.0 CTA); and (c) the BALB/c 3T3 CTA.</p> <p>ECVAM coordinated an international study that was designed to address issues of CTA protocols standardisation, transferability and reproducibility in three protocol variants for the SHE CTA (at pH 6.7 and pH 7.0) and the BALB/c 3T3 assay. The study of the three test methods was peer reviewed by the ESAC that issued a peer review report and an ESAC opinion. Based on the available documents (study reports, ESAC opinion, OECD DRP), EURL ECVAM recommended that an OECD TG be developed for the SHE CTA and that further investigations were needed to confirm the performance of the BALB/c 3T3 CTA.</p> <p>Moreover, the recommendation stated that in conjunction with other available data, the CTAs have the potential of partial replacement or reduction when used in a weight of evidence approach for hazard identification and risk assessment.</p>
<p>EURL ECVAM (2013) Recommendation on the Cell Transformation Assay based on the Bhas 42 cell line</p>	<p>CTA in Bhas 42 cell line that offers a relatively higher through-put than other CTA variants.</p>

http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-cell-transformation-assay-based-on-the-bhas-42-cell-line	
Regulatory Requirements for Carcinogens across Sectors	
<p>European Commission, 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353.</p>	<p>Carcinogenicity (Section 3.6)</p> <p>CATEGORY 1 Known or presumed human carcinogens A substance is classified in <i>Category 1</i> for carcinogenicity on the basis of epidemiological and/or animal data. A substance may be further distinguished as:</p> <p style="padding-left: 40px;"><u>Category 1A</u>, known to have carcinogenic potential for humans, classification is largely based on human evidence, or</p> <p style="padding-left: 40px;"><u>Category 1B</u>, presumed to have carcinogenic potential for humans, classification is largely based on animal evidence.</p> <p>The classification in Category 1A and 1B is based on strength of evidence together with additional considerations. Such evidence may be derived from:</p> <ul style="list-style-type: none"> • human studies that establish a causal relationship between human exposure to a substance and the development of cancer (known human carcinogen); or • animal experiments for which there is sufficient evidence to demonstrate animal carcinogenicity (presumed human carcinogen). <p>In addition, on a case-by-case basis, scientific judgment may warrant a decision of presumed human carcinogenicity derived from studies showing limited evidence of carcinogenicity in humans together with limited evidence of carcinogenicity in experimental animals</p> <p>CATEGORY 2 Suspected human carcinogens The placing of a substance in Category 2 is done on the basis of evidence obtained from human and/or animal studies, but which is not sufficiently convincing to place the substance in Category 1A or 1B, based on strength of evidence together with additional considerations. Such evidence may be derived either from limited evidence of carcinogenicity in human studies or from limited evidence of carcinogenicity in animal studies.</p>
<p>GHS (2013) UN (2013) Globally Harmonized System of Classification and Labelling of Chemicals (GHS), 5th revised edition. United Nations, New York and Geneva, 2013 http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf.</p>	<p>Carcinogenicity (Chapter 3.6)</p> <p>3.6.1 Definitions The term carcinogen denotes a substance or a mixture which induces cancer or increases its incidence. Substances and mixture which have induced benign and malignant tumours in well performed experimental studies on animal are considered also to be presumed or suspected human</p>

	<p>carcinogens unless there is strong evidence that the mechanism of tumour formation is not relevant for humans.(...)</p> <p><i>3.6.2 Classification criteria for substances</i></p> <p>CATEGORY 1 Known or presumed human carcinogens</p> <p>Category 1A: Known to have carcinogenic potential for humans; the placing of a substance is largely based on human evidence</p> <p>Category 1B: Presumed to have carcinogenic potential for humans; the placing of a substance is largely based on animal evidence.</p> <p>CATEGORY 2 Suspected human carcinogens</p> <p>Based on evidence obtained from human and/or animal but not sufficiently convincing to place a substance in Category 1.</p>										
<p>IARC (2006) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. WHO, International Agency for Research on Cancer, volumes 1-107</p>	<p>Evaluation and Rational (Preamble, par. 6)</p> <table border="0"> <tr> <td>Group 1</td> <td>Carcinogenic to humans</td> </tr> <tr> <td>Group 2A</td> <td>Probably carcinogenic to humans</td> </tr> <tr> <td>Group 2B</td> <td>Possibly carcinogenic to humans</td> </tr> <tr> <td>Group 3</td> <td>Not classifiable as to its carcinogenicity to humans</td> </tr> <tr> <td>Group 4</td> <td>Probably not carcinogenic to humans</td> </tr> </table>	Group 1	Carcinogenic to humans	Group 2A	Probably carcinogenic to humans	Group 2B	Possibly carcinogenic to humans	Group 3	Not classifiable as to its carcinogenicity to humans	Group 4	Probably not carcinogenic to humans
Group 1	Carcinogenic to humans										
Group 2A	Probably carcinogenic to humans										
Group 2B	Possibly carcinogenic to humans										
Group 3	Not classifiable as to its carcinogenicity to humans										
Group 4	Probably not carcinogenic to humans										
<p>ICH Guideline (1995) S1A ICH harmonized Tripartite Guideline, Guideline on the Need for Carcinogenicity Studies of Pharmaceuticals, S1A, Current Step version 4, November 1995.</p> <p>ICH Guideline (2009) S1B Guideline on carcinogenicity testing of pharmaceuticals. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002735.pdf.</p> <p>ICH Guideline (2008) S1C(R2) Dose selection for carcinogenicity Studies of Pharmaceuticals, S1C (R2), Current Step version 4, 11 March 2008.</p> <p>ICH Guideline (2009) M3 (R2)</p>	<p><i>Pharmaceuticals:</i> (S1A) The guideline sets the factors in need to be considered for carcinogenicity testing: exposure time, cause of concern, genotoxicity, indication and patient population, route of exposure, extent of systemic exposure, endogenous peptides, and analogues (Paragraphs 4.1 to 4.7).</p> <p>Carcinogenicity studies <u>should be performed for any pharmaceutical</u> whose expected clinical use is continuous for at least 6 months.</p> <p>Also for pharmaceuticals used frequently in an intermittent manner in the treatment of chronic or recurrent conditions the studies are generally needed.</p> <p>Studies are also recommended:</p> <ul style="list-style-type: none"> -for some drugs if there is concern about their carcinogenic potential, derived from previous carcinogenicity studies or other toxicity tests or QSARs predictions. -for pharmaceuticals applied topically, unless a proof of poor systemic exposure is shown. 										

<p>Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. June 2009.</p> <p><i>Concept Paper (2012)</i> S1: Rodent Carcinogenicity Studies for Human Pharmaceuticals, Dated and endorsed by the Steering Committee on 14 November 2012</p> <p><i>ICH Guideline S1 (2012)</i> Regulatory notice on changes to core guideline on rodent carcinogenicity testing of pharmaceuticals. EMA/CHMP/ ICH/752486/2012 December 2012.</p>	<p>Unequivocally genotoxic compounds, instead, need not to be subjected to long-term carcinogenicity studies. The detection of early tumorigenic effects may be necessary if the drug is intended for chronic treatment (up to one year).</p> <p>Endogenous peptides or proteins and their analogues, produced by chemical synthesis, by extraction/purification from animal/human source or biotechnological methods may require special considerations.</p> <p>Not generally needed but, conduct of long-term carcinogenicity studies in rodent species are required for: 1) products where there are significant differences in biological effects with the natural counterparts; 2) for products where modifications lead to significant changes in structure; 3)for products resulting in humans in a significant increase over the existing local or systemic concentration.</p> <p>(S1B) The guideline embraces all the pharmaceutical agents that need carcinogenicity testing (S1A) apart from biotechnology-derived pharmaceuticals which are considered in a specific guideline (S6). Experimental approaches as from Paragraph 4.2: The basic scheme is: -One long-term rodent carcinogenicity study, generally rat, plus -One other study which provides additional information such as: a) short or medium-term <i>in vivo</i> rodent test systems, either or b) long-term study in a second rodent species (mouse generally)</p> <p>Mechanistic studies are considered for the interpretation of the results (i.e. tumor findings) from carcinogenicity studies: the above include cellular changes and biochemical measurements.</p> <p>Additional genotoxicity studies are also considered; they might be invoked for compounds negative in the standard test battery but showing effects in carcinogenicity studies via non-clear epigenetic mechanisms.</p> <p>S1C (R2) This document addresses the criteria for the selection of the high dose to be used in carcinogenicity studies on new therapeutic agents to harmonise current practices and improve the design of studies. In this second revision, the pharmacokinetic endpoint of 25 is declared to be applicable also for pharmaceuticals with positive genotoxicity signals. This change has implications on "Refinement" (one of the 3R's) in enhancing the welfare, i.e., reducing the pain or discomfort of the animals at the maximally tolerated dose (MTD).</p>
--	---

	<p>M3 (R2) To recommend international standards for, and promote harmonisation of, the nonclinical safety studies to support human clinical trials of a given scope and duration, and for the marketing authorization of drug products</p> <p>Concept Paper A change to the current S1 harmonised Guidelines on rodent carcinogenicity testing is proposed to introduce a more comprehensive and integrated approach to addressing the risk of human carcinogenicity of pharmaceuticals, clarify and update, without compromising safety, the criteria for deciding whether the conduct of a two-year rodent carcinogenicity study of a given pharmaceutical would add value to this risk assessment. In November 2012, the SC endorsed the revision of both the S1 Concept Paper and Business Plan to provide clarification concerning how the prospective data gathering period should be integrated in the normal ICH Step process. The revised S1 Concept Paper and Business Plan describe the S1 strategy which consists of a draft "Regulatory Notice for Public Input" which would be issued by each ICH regulatory health authority to solicit comments from the public to the proposal, the procedure, and the specific weight-of-evidence criteria. A final "Regulatory Notice" is planned to be published in June 2014 and will mark the beginning of the prospective data (CADs: Carcinogenicity Assessment Documents) collection period. After collecting and incorporating results from the prospective analyses, a Step 2 document is planned to be published in November 2016, and a Step 4 document finalised in November 2017.</p>
<p>VICH GL28 (SAFETY: CARCINOGENICITY) (2005) <i>Studies to evaluate the safety of residues of veterinary drugs in human food: Carcinogenicity Testing.</i> February 2005 For implementation at Step 7 – Final (Chapter 2.)</p>	<p><i>Veterinary Drugs:</i> This guideline sets out a data-driven decision pathway to determine the need to conduct carcinogenicity studies. It also provides guidance on the conduct of carcinogenicity studies.</p> <p>The decision to undertake carcinogenicity testing should take into consideration:</p> <ol style="list-style-type: none"> 1) the results of genotoxicity tests, 2) structure-activity relationships, 3) findings from systemic toxicity tests that may be relevant to neoplasia in longer term studies. It should also take into consideration any known species specificity of the mechanism of toxicity. <p><i>Any differences in metabolism between the test species, target animal species, and human beings should be taken into consideration.</i></p> <p>Genotoxic and non-genotoxic carcinogens:</p>

	<p>Clearly negative results for genotoxicity will usually be taken as sufficient evidence of a lack of carcinogenic potential via a genotoxic mechanism.</p> <p>Non-genotoxic compounds do not need to be routinely tested for carcinogenicity. Such tests may however be required if, for example:</p> <ol style="list-style-type: none"> 1) the compound is a member of a chemical class known to be animal or human carcinogens, 2) available systemic toxicity studies with the compound identify potentially preneoplastic lesions or findings indicative of neoplasia, or 3) systemic toxicity studies indicate that the compound may be associated with effects known to be linked with epigenetic mechanisms of carcinogenicity that are relevant to humans. <p><i>In vivo</i> testing OECD TGs: 451 and 453.</p> <p>Carcinogenicity bioassays consisting of a two-year rat study and an 18-month mouse study are generally required. With appropriate scientific justification, carcinogenicity studies may be carried out in one rodent species, preferably the rat. A positive response in either test species will be considered indicative of carcinogenic potential.</p>
<p>European Commission, 2006. Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Communities L396, 1-849.</p> <p>And</p> <p>ECHA (2012) Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific Guidance Version 2.0 November 2012, R7.7. Mutagenicity and Carcinogenicity.</p>	<p><i>Chemicals:</i></p> <p>Proposals for conducting a carcinogenicity test should be made with regard to the potential risk to human health and with consideration of the actual or intended production and/or use pattern.</p> <p>REACH only specifies a carcinogenicity test for substances at the Annex X tonnage level (≥ 1000 t/y).</p> <p>Annex X (1000 tpa-more) par. 8.9.1</p> <p>A carcinogenicity study may be proposed by the registrant or may be required by the Agency in accordance with Articles 40 or 41 if:</p> <ul style="list-style-type: none"> • the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure; and • the substance is classified as mutagen category 3 or there is evidence from the repeated dose study (ies) that the substance is able to induce hyperplasia and/or pre-neoplastic lesions. <p>If the substance is classified as mutagen category 1 or 2, the default presumption would be that a genotoxic mechanism for carcinogenicity is likely. In these cases, a carcinogenicity test will normally not be required.</p>

	<p><i>However, REACH also requires that carcinogenic substances at all tonnage levels be identified as substances of high concern, taking into account information from all available relevant sources.</i></p> <p>At the tonnage levels below 1000 t/y, the main concern is for those chemicals that are genotoxic. Chemicals may cause cancer secondary to other forms of toxicity, but protection of human health against the underlying toxicity (e.g., as identified from a repeat-dose toxicity study) will also protect against cancer that is secondary to that toxicity. It is noted, though, that some of these non-genotoxic carcinogens, when not classified for any other property and not identified as such in (limited) repeated dose toxicity studies will go unidentified; this also regards the risks associated with human exposures.</p> <p><u>Testing strategy for carcinogenicity</u></p> <p>As for other endpoints, the following three steps apply for the assessment of carcinogenicity (i.e. the hazard, underlying mode of action, and potency) for substances at each of the tonnage levels specified in Annexes VII to X of REACH.</p> <ol style="list-style-type: none"> i. Gather and assess all available test and non-test data from read-across/proper chemical category and suitable predictive models. Examine the Weight of Evidence that relates to carcinogenicity. ii. Consider whether the standard information requirements are met. iii. Ensure that the information requirements of Annexes VII and VIII are met; make proposals to conform to Annexes IX and X.
<p>European Commission, 2009a. Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Communities L342, 59-209.</p> <p>And</p> <p>SCCS'S Notes of Guidance for the testing of cosmetics substances and their safety evaluation, 8th Revision, SCCS/1501/12 (par 3-4.8).</p>	<p><i>Cosmetics:</i></p> <p>Before the testing/marketing ban of the 7th Amendment of the Cosmetics Directive [2003/15/EC] on cosmetic ingredients, the most commonly performed carcinogenicity tests were:</p> <ol style="list-style-type: none"> i. the Carcinogenicity test [EC B.32, OECD 451] or ii. the combined chronic toxicity / carcinogenicity test [EC B.33, OECD 453]. <p>Under this testing and marketing ban, <i>in vivo</i> testing to investigate the carcinogenic potential of substances is no longer possible.</p> <p>Unfortunately, at present no validated methods to study carcinogenicity are available.</p> <p>The <i>in vitro</i> <u>Cell Transformation Assay</u> (CTA) is at a late stage of development.</p>

	<p>Without the 2-year bioassay, it is very difficult if not impossible to conclude on the carcinogenicity of substances. As far as genotoxic compounds are concerned, <i>in vitro</i> mutagenicity tests are quite well developed. Due to the relation between mutations and cancer, these genotoxicity tests can be seen as a pre-screen for cancer. A positive result in one of the genotoxicity tests may be indicative enough to consider a compound as putatively carcinogenic. In combination with the CTA, this indication may even be stronger. However, as carcinogenicity is a multi-hit/multi step process, it can (for the time being) not be mimicked by <i>in vitro</i> tests. Today, any reliable, justified conclusion on the carcinogenicity of a substance needs <i>in vivo</i> tests.</p> <p>The situation is different for the non-genotoxic carcinogens.</p> <p>Before the animal testing and marketing ban, non-genotoxic carcinogens were detected by the (sub) chronic repeated dose studies, including the carcinogenicity test. Alternatives for these <i>in vivo</i> tests to detect non-genotoxic carcinogens, however, are not available with the exception of the CTA. Therefore, currently, <i>in vivo</i> rodent studies are essential to detect non-genotoxic substances. An extensive review of the actual status of <i>in vitro</i> carcinogenicity testing can be found in a JRC report [Adler et al. 2011].</p>
<p>European Commission, 2009b. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-47.</p> <p>Commission Regulation (EU), 2013, No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.</p> <p>Commission Regulation (EU), 2013, No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council</p>	<p><i>Plant Protection Products:</i> Long-term toxicity and carcinogenicity</p> <p>The results of the long-term studies conducted and reported, taken together with other relevant data and information on the active substance, shall be sufficient to permit the identification of effects, following repeated exposure to the active substance, and in particular shall be sufficient to:</p> <ol style="list-style-type: none"> i. identify adverse effects resulting from long-term exposure to the active substance, ii. identify target organs, where relevant, iii. establish the dose-response relationship, iv. establish the NOAEL and, if necessary, other appropriate reference points. <p>Correspondingly, the results of the carcinogenicity studies taken together with other relevant data and information on the active substance, shall be sufficient to permit the evaluation of hazards for humans, following repeated exposure to the active substance, and in particular shall be sufficient:</p> <ol style="list-style-type: none"> (a) to identify carcinogenic effects resulting from long-term exposure to the active substance; (b) to establish the species, sex, and organ specificity of tumours induced; (c) to establish the dose-response relationship;

<p>concerning the placing of plant protection products on the market (Text with EEA relevance)</p>	<p>(d) where possible, to identify the maximum dose eliciting no carcinogenic effect; (e) where possible, to determine the mode of action and human relevance of any identified carcinogenic response.</p> <p>Circumstances in which required</p> <p>The long-term toxicity and carcinogenicity of all active substances shall be determined. If in exceptional circumstances it is claimed that such testing is unnecessary, that claim shall be fully justified.</p> <p>Test conditions</p> <p>A long-term oral toxicity study and a long-term carcinogenicity study (two years) of the active substance shall be conducted using rat as test species; where possible these studies shall be combined.</p> <p>A second carcinogenicity study of the active substance shall be conducted using mouse as test species, unless it can be scientifically justified that this is not necessary. In such cases, scientifically validated alternative carcinogenicity models may be used instead of a second carcinogenicity study. If comparative metabolism data indicate that either rat or mouse is an inappropriate model for human cancer risk assessment, an alternative species shall be considered. Experimental data, including the elucidation of the possible mode of action involved and relevance to humans, shall be provided where the mode of action for carcinogenicity is considered to be non-genotoxic.</p> <p>Where submitted, historical control data shall be from the same species and strain, maintained under similar conditions in the same laboratory and shall be from contemporaneous studies. Additional historical control data from other laboratories may be reported separately as supplementary information.</p>
<p>European Union Regulation, 2012. Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L 167, 1-116.</p>	<p><i>Biocides:</i> Annex II, Data Requirement</p> <p><i>8.11 Carcinogenicity study</i></p> <ul style="list-style-type: none"> • A carcinogenicity study does not need to be conducted if: the substance is classified as mutagen category 1A or 1B. The default presumption would be that a genotoxic mechanism for carcinogenicity is likely. In these cases, a carcinogenicity test will normally not be required.

	<p><i>8.11.1. Combined carcinogenicity study and long-term repeated dose toxicity</i></p> <ul style="list-style-type: none"> • Rat, oral route of administration is the preferred route. If an alternative route is proposed a justification must be provided. • For evaluation of consumer safety of active substances that may end up in food or feed, it is necessary to conduct toxicity studies by the oral route <p><i>8.11.2 Carcinogenicity testing in a second species</i></p> <ul style="list-style-type: none"> • A second carcinogenicity study should normally be conducted using the mouse as test species. • For evaluation of consumer safety of active substances that may end up in food or feed, it is necessary to conduct toxicity studies by the oral route
Guidance Documents, Opinions and Reviews on Carcinogenicity testing	
US EPA (2005) Guidelines for Carcinogen Risk Assessment US EPA EPA/630/P-03/001F, March 2005	This is a revision and replacement of the U.S. Environmental Protection Agency's (EPA's, or the Agency's) Guidelines for Carcinogen Risk Assessment, published in 51 FR 33992, September 24, 1986 (U.S. EPA, 1986a) and the 1999 interim final guidelines (U.S. EPA, 1999a; see U.S. EPA 2001b). The documents provide guidance for developing and using risk assessments. They also provide basic information to the public about the EPA Agency's risk assessment methods.
OECD Guidance Document (2012a), no. 116 On the conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Support test Guidelines 451, 452 AND 453 2ND EDITION, ENV/JM/MONO (2011) 47; 13-Apr-2012	This guidance provides additional information on the conduct of studies performed using TG 451, 452 and TG 453. Its objective is to assist users of the TGs to select the most appropriate methodology to assess the chronic toxicity and carcinogenicity of a test chemical so that particular data requirements can be met while reducing animal usage if possible/appropriate.
OECD (2010d) OECD principles for the validation, for regulatory purposes, of (quantitative) structure-activity relationship models. http://www.oecd.org/dataoecd/33/37/37849783.pdf	The document describes the principles to facilitate the consideration of a (Q)SAR model for regulatory purposes.
SCCS'S (2012) Notes of Guidance for the testing of cosmetics substances and their safety evaluation, 8th Revision, SCCS/1501/12.	This document contains recommendations for the testing of cosmetics, in light of the full bun of <i>in vivo</i> experiments (March 13, 2013).
ECHA (2011), Guidance on information requirements and chemical safety assessment Part B: Hazard assessment. Version 2.1 December 2011.	This document contains guidance on REACH explaining the REACH obligations and how to fulfil them: this part refers to the Hazard Assessment.
ECHA Guidance (2012). Guidance on information	This document contains guidance on REACH explaining the REACH obligations and how to fulfil

requirements and chemical safety assessment Chapter R.7a: Endpoint specific Guidance Version 2.0 November 2012, R7.7. Mutagenicity and Carcinogenicity.	them: this part refers to the Endpoints of Mutagenicity and Carcinogenicity.
SCCP (2008). Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products SCCP/1171/08 (SCCS/SCHER/SCENIHR adopted on June 8, 2012).	The SCCP (now SCCS)/SCHER/SCENIHR (SCs) evaluated potential applications of the Threshold of Toxicological Concern (TTC) approach for human health risk assessment of chemical substances. The opinion focused on the potential applications of the TTC concept for cosmetics and other consumer products in relation to the mandates of the three SCs. The application of TTC in other areas, such as food, pharmaceuticals or EU chemical legislation (REACH) was not assessed, although such applications were described for completeness.
EMA (2006) European Medicines Evaluation Agency. Committee for Medicinal Products for Human Use (CHMP). Guideline on the limits of genotoxic impurities. CPMP/SWP/5199/02. London, 28 June 2006. EMEA/CHMP/QWP/ 251344/2006. http://www.emea.europa.eu/pdfs/human/swp/519902en.pdf	This Guideline describes a general framework and practical approaches, mainly establishment of TTC, on how to deal with genotoxic impurities in new active substances.
EMA (2008) Committee for Medicinal Products for Human Use (CHMP). Questions & Answers on the CHMP Guideline on the limits of Genotoxic Impurities. Rev 1. London, 26 June 2008. Doc.Ref. EMEA/CHMP/SWP/431994/2007	The aim of this question-and-answer document was to provide clarification and harmonisation of the 'Guideline on the limits of genotoxic impurities' (EMA/CHMP/QWP/251344/2006), published in 2006. When and how applying TTC or QSAR approaches.
ECHA (2010a) Practical guide on how to report (Q)SARs. http://echa.europa.eu/doc/publications/practical_guides/pg_report_qsars.pdf , ECHA-10-B-10-EN	This practical guide provides an overview of important aspects when predicting properties of substances using (Q)SAR models as defined in the REACH Regulation.
ECHA (2010b) Practical guide on how to report categories and read-across. ECHA-10-B-11-EN http://echa.europa.eu/doc/publications/practical_guides/pg_report_readacross_categ.pdf	This practical guide provides an overview of important practical aspects on how to develop and report in IUCLID 5 read-across and/or a chemical category approach for substances to be registered under the REACH Regulation.
JECFA (1996) A procedure for the safety evaluation of flavouring agents. WHO Food Additive Series 35. Annex 5. World Health Organization, Geneva	TTC approach application for the safety evaluation of flavouring agents. The objective of this paper was to provide a procedure that can be used by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) for the safety evaluation of flavouring substances.
JECFA (1997) Safety evaluation of certain food additives and contaminants. Forty-ninth Report of the Joint FAO/WHO	Six groups of flavouring agents were evaluated using the Procedure for the Safety Evaluation of Flavouring Agents, taking into account their metabolism.

Expert Committee on Food Additives. WHO Food Additive Series 40. World Health Organization, Geneva	
EFSA (2004) European Food Safety Authority. Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food. Opinion on Flavouring Group FGE.03 Acetals of branched- and straight-chain aliphatic saturated primary alcohols and branched- and straight-chain saturated aldehydes, and an orthoester of formic acid, from chemical groups 1 and 2. Opinion expressed on 7 October 2004. EFSA J. 107, 1-59.	The Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food was asked to evaluate 42 flavouring substances in the Flavouring Group Evaluation FGE.03.
Peer Consultation on Health Canada Draft Weight of Evidence Framework for Genotoxic Carcinogenicity (2005) http://www.tera.org/peer/comhaz/HC%20QSAR%20Peer%20Consultation%20Overview.pdf	Evaluation of the available methods for the prediction of carcinogenic potential.
Adler, S., Basketter, D., Creton, S., et al., (2011) Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. <i>Archives of Toxicology</i> 85 (5): 367-485.	<p>In regard of the 7th Amendment to the EU Cosmetics Directive to prohibit animal-tested cosmetics on the market starting 2013, the European Commission invited stakeholder bodies (industry, nongovernmental organisations, EU Member States, and the Commission's Scientific Committee on Consumer Safety) to identify scientific experts in five toxicological areas, i.e. toxicokinetics, repeated dose toxicity, carcinogenicity, skin sensitisation, and reproductive toxicity for which the Directive foresees that the 2013 deadline could be further extended in case alternative and validated methods would not be available in time.</p> <p>The experts evidenced impediments for a full replacement of animal testing, on due time, relating the carcinogenicity toxicological area.</p> <p>The carcinogenic potential of a chemical substance is far to be fully determined and/or mimicked by the use of non-animal testing because of the complexity of the carcinogenesis process, the multi-stages type of its evolution and complex biological interactions. The 'Gold Standard' for cancer hazard evaluation is the 2-year cancer assay in rodents however, the animal testing ban under the 7th Amendment to the Cosmetics Directive will have a strong impact on the ability to conduct quality risk assessment for carcinogenic potential. The authors underlined that the impact is due not only to the ban of the cancer assay but mainly to the ban of <i>in vivo</i> genotoxicity testing, any repeated dose toxicity testing and others such as <i>in vivo</i> toxicokinetics studies and <i>in vivo</i> mechanistic assays.</p> <p>Beyond the availability of standard <i>in vitro</i> genotoxicity tests and several <i>in vitro</i> short-term tests (many of which at different stages of regulatory validation process), no sufficient <i>in vitro</i> test do exist to full replacement of animal testing. Moreover, the available tests are focused on hazard</p>

	evaluation and do not support a full safety assessment.
Gottmann E, Kramer S, Pfahringer B & Helma C (2001). Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. <i>Environmental Health Perspectives</i> 109 (5): 509-14.	In this study, 121 replicate rodent carcinogenicity assays from National Cancer Institute/National Toxicology Program and literature and the Carcinogenic Potency Database (CPDB) were compared in order to estimate the reliability of these experiments. A concordance of 57% between the overall rodent carcinogenicity classifications from both sources was found. The results indicate that rodent carcinogenicity assays were much less reproducible than previously expected.
Knight A, Jonker MJ & Bruning J (2005). Which drugs cause cancer? <i>British Medical Journal USA</i> 5: 477-478.	The authors highlighted the extremely low human carcinogenicity predictivity of <i>in vivo</i> experiments. The burden of cancer disease in United States is way too high despite the billions dollars spent in Research and millions of laboratory animals sacrificed, since the 'War against Cancer' was launched in the early '70s. The problem with <i>in vivo</i> tests is not the lack of sensitivity but rather the lack of human specificity. The adoption of alternatives is recommended such as QSAR models, enhanced <i>in vitro</i> assays and cDNA microarrays, the potential of which is to yield superior human specificity results.
Knight A, Bailey J & Balcombe J (2006). Animal carcinogenicity studies: Implications for the REACH system. <i>Alternatives to Lab Animals</i> 34, Suppl. 1: 139-147.	Based on a survey of the US Environmental Protection Agency's (EPAs) toxic chemicals database revealing that, for a majority of the chemicals of greatest public health concern (58.1%), animal carcinogenicity data are inadequate to support classifications of probable human carcinogen or non-carcinogen, and show several significant differences with IARC assessments of identical chemicals, the authors propose the replacement of animal carcinogenicity bioassays with a tiered combination of non-animal assays. The alternative strategy is expected to yield a weight-of-evidence characterisation of carcinogenic risk with superior human predictivity. Additional advantages include substantial savings of financial, human and animal resources, and potentially greater insights into mechanisms of carcinogenicity.
Alden CL, Lynn A, Bourdeau A, Morton D, Sistare FD, Kadambi VJ & Silverman L (2011). A critical review of the effectiveness of rodent pharmaceutical carcinogenesis testing in predicting for human risk. <i>Veterinary Pathology</i> 48 (3): 772-84.	Reviewing the status of carcinogenicity testing in the pharmaceutical sector, the authors proposed the appropriate use of acute, subchronic, chronic, and special toxicology tests to identify the major associated cancer risk factors, specifically, hormonal modulation, immunosuppression, genetic toxicity, and chronic toxicity. They pointed at the existence of significant opportunities already available for improving the effectiveness and efficiency of the current cancer risk assessment paradigm.
Luijten, M, Muller JJA, Hernández LG, van der Ven LTM & van Benthem J (2012). Prediction of carcinogenic potential of substances using repeated dose toxicity data. RIVM Report	The group, through an extensive literature review, proposes as an alternative approach to determine the carcinogenic features of substances the use of data from sub-chronic repeated dose toxicity studies. This approach could lead to a substantial reduction in the number of carcinogenicity studies

340700006, pp 1-18.	performed without compromising human safety.
Paules RS, Aubrecht J, Corvi R, Garthoff B & Kleinjans JC (2011). Moving Forward in Human Cancer Risk Assessment. Environmental Health Perspectives 119:439-443.	The review focuses on the state-of-the-art in developing alternative testing strategies for carcinogenicity, emphasising on potential contributions from omics technologies.
Maurici D, Aardema M, Corvi R, Kleber M, Krul C, Laurent C., Loprieno N, Pasanen M, Pfuhler S, Phillips B, Sabbioni E, Sanner T & Vanparys P (2005). Genotoxicity and mutagenicity. Alternatives to Lab Animals 33 Suppl 1: 117-130.	The aim of this report was to provide an objective state of play of the current status of alternative methods/strategies and the prospects for their validation and regulatory acceptance so that they could be used for replacing animal tests in the safety assessment of cosmetic products as required by the EU Cosmetics Directive. Regarding the endpoint genotoxicity/mutagenicity, the experts were of the opinion that a total replacement of animal testing was not feasible in the short-term and this would depend, besides the development of <i>in vitro</i> tests on skin models, also on the progress in the fields of toxicokinetics and toxicogenomics.
Benigni R. (2012). Alternatives to the carcinogenicity bioassay for toxicity prediction: are we there yet? Expert Opinion Drug Metabolism & Toxicology 8 (4): 407-17.	The expert in this review summarizes the theories on the early steps of carcinogenesis and reports about alternative detection methods for carcinogens based on genetic toxicology, structure--activity relationships and cell transformation assays.
Benfenati E, Benigni R, DeMarini DM, Helma C, Kirkland D, Martin TM, Mazzatorta P, Ouédraogo-Arras G, Richard AM, Schilter B, Schoonen WGEJ, Snyder RD & Yang C (2009). Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives. Journal of Environmental Science and Health Part C Environmental Carcinogenesis & Ecotoxicology Reviews 27 (2): 57-90.	The authors reports on the advances of new technologies and emerging methodologies, software tools, QSAR models and <i>in silico</i> methods for predicting genotoxicity and carcinogenicity. The use of these methodologies, particularly high-throughput assays and data mining in combination with <i>in silico</i> methods, is encouraged for industry and regulators assistance and ultimately for public health protection.
Hereceg Z, Lambert M-P, van Veldhoven K, Demetriou C, Vineis P, Smith MT, Straif K & Wild CP (2013). Towards incorporating epigenetic mechanisms into carcinogen identification and evaluation. Carcinogenesis doi:10.1093/carcin/btg212. 1-13.	The authors, taking into account the remarkable progress in the field of epigenetics, summarize the current status of incorporation of epigenetic mechanisms into carcinogen evaluation by the IARC.
Mechanistic Understanding of Carcinogenesis: Genotoxic and non-genotoxic driven mechanisms	
Fowler P, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S & Carmichael P (2012). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I.	As part of an European Cosmetics Industry Association initiative for improvement of <i>in vitro</i> mammalian cell assays, Fowler and co-workers have compared several rodent cell lines (V79, CHL, CHO) with p53-competent human peripheral blood lymphocytes (HuLy), TK6 human

<p>Choice of cell type. Mutation Research 742 (1-2): 11-25.</p>	<p>lymphoblastoid cells, and the human liver cell line, HepG2 since, p53-deficiency in many of the rodent cell lines has been considered a key factor in the poor testing predictivity. The authors have compared <i>in vitro</i> micronucleus (MN) induction following treatment with 19 compounds that were accepted as producing misleading or "false" positive results in <i>in vitro</i> mammalian cell assays. The rodent cell lines (V79, CHO and CHL) resulted consistently more susceptible to cytotoxicity and MN induction than p53-competent cells, and were therefore more susceptible to giving misleading positive results. The authors suggested that a reduction in the frequency of misleading positive results could be achieved by careful selection of the mammalian cell type for genotoxicity testing.</p>
<p>Fowler P, Smith R, Smith K, Young J, Jeffrey L, Kirkland D, Pfuher S & Carmichael P (2012). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. Mutation Research 747 (1): 104-17.</p>	<p>In line with the outcomes of the analysis mentioned in the above paper (Fowler et al., 2012a); the authors investigated the impact of different toxicity measures, commonly used <i>in vitro</i> cytogenetic assays, on the occurrence of misleading positive results. The data showed that estimating toxicity by relative cell count (RCC) or replication index (RI) consistently underestimated the toxicity observed by other measures (Relative Population Doubling, RPD, or Relative Increase in Cell Count, RICC). RCC and RI were more likely to lead to selection of concentrations for micronucleus scoring that were highly cytotoxic and thus could potentially lead to artefacts of toxicity being scored (elevated levels of apoptosis and necrosis), generating misleading positive results. The authors suggested that a further reduction in the frequency of misleading positive results in <i>in vitro</i> cytogenetic assays could be achieved (with 19 chemicals clearly giving misleading results) by avoiding the use of toxicity measures that underestimate the level of toxicity induced.</p>
<p>Kirkland D, Aardema M, Henderson L & Müller L (2005). Evaluation of the ability of a battery of three <i>in vitro</i> genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. Mutation Research 584: 1-256.</p>	<p>This paper evaluated how the currently popular <i>in vitro</i> genotoxicity tests – Ames test, MLA and a test for clastogenicity (<i>in vitro</i> micronucleus or chromosomal aberration test) – performed in their ability to discriminate rodent carcinogens and non carcinogens. The work deals separately with the performance indicators (sensitivity first, then specificity) of the individual tests and various combinations of two or three tests, and proposes reasons why some rodent carcinogens are not detected in this battery. It also proposes a new method to look at the balance of rising sensitivity accompanied by falling specificity to help choose the best individual or combinations of tests.</p>
<p>Kirkland D, Pfuher S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier J-R, Müller L, Nohynek, G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P & White P (2007). How to reduce false positive results when</p>	<p>A workshop organised by ECVAM in 2006 on "How to reduce the false positive results when undertaking <i>in vitro</i> genotoxicity testing and thus avoid unnecessary follow-up animal tests" addressed (i) whether it is possible to choose existing cell systems which give lower rates of false positive results, (ii) whether modifications of existing protocols and cell systems may result in lower false positive results, (iii) the performance of new test systems showing promise of improved specificity and definition of needs for the development of new tests.</p>

<p>undertaking <i>in vitro</i> genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. Mutation Research 628: 31-55.</p>	
<p>Widschwendter M & Jones PA (2002). DNA methylation and breast carcinogenesis. Oncogene 21: 5462-5482.</p>	<p>Knowledge about breast carcinogenesis has accumulated during the last decades but has barely been translated into strategies for early detection or prevention of this common disease. Changes in DNA methylation have been recognized as one of the most common molecular alterations in human neoplasia and hypermethylation of gene-promoter regions is being revealed as one of the most frequent mechanisms of loss of gene function. The heritability of methylation states and the secondary nature of the decision to attract or exclude methylation support the idea that DNA methylation is adapted for a specific cellular memory. There are six novel capabilities a cell has to acquire to become a cancer cell: limitless replicative potential, self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death, sustained angiogenesis and tissue invasion and metastasis. This review highlights how DNA-methylation contributes to these features and offers suggestions about how these changes could be prevented, reverted or used as a 'tag' for early detection of breast cancer or, preferably, for detection of premalignant changes.</p>
<p>Baylin SB & Ohm JE (2006). Epigenetic gene silencing in cancer – a mechanism for early oncogenic pathway addition? Nature Reviews Cancer 6: 107-116.</p>	<p>This review focuses on epigenetic changes, i.e. chromatin alterations and DNA methylation which associate with all stages of tumour formation and progression. DNA methylation particularly occurs at promoter regions of genes that regulate important cell functions. Recent evidence indicates that epigenetic changes might 'addict' cancer cells to altered signal-transduction pathways during the early stages of tumour development. Dependence on these pathways for cell proliferation or survival allows them to acquire genetic mutations in the same pathways, providing the cell with selective advantages that promote tumour progression. The authors highlight the relevance of strategies to reverse epigenetic gene silencing in cancer prevention and therapy.</p>
<p>Esteller M (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nature Reviews Genetics 8: 286-298.</p>	<p>The author reviews several studies exploring the mosaic patterns of DNA methylation and histone modification in cancer cells on a 'gene-by-gene basis': many of those reporting the role of CpG-island-promoter hypermethylation in the transcriptional silencing of tumour-suppressor genes. However, the author promotes the 'genome-wide approach' to provide both biological insight and new avenues for translational research.</p>
<p>Shanks N, Greek R & Greek J (2009). Are animal models predictive for humans? Philosophy, Ethics & Humanities in Medicine 4: 1-20.</p>	<p>The focus of this essay is the scientific term predict and whether there is credible evidence that animal models, especially in toxicology and pathophysiology, can be used to predict human outcomes. Indeed, Shanks and co-workers state that animal models fall far short of being able to predict human responses to drugs and other chemicals.</p>

<p>Hattis D, Chu M, Rahmioglu N, Goble R, Verma, P, Hartman K & Kozlak M (2009). A preliminary operational classification system for nonmutagenic modes of action for carcinogenesis. <i>Critical Reviews in Toxicology</i> 39 (2): 97-138.</p>	<p>The article proposes a system of categories for non-mutagenic modes of action for carcinogenesis. The classification is of modes of action rather than individual carcinogens, because the same compound can affect carcinogenesis in more than one way. Basically, the authors categorize modes of action as: (1) co-initiation; (2) promotion; (3) progression and (4) multiphase. Agents that act at relatively early stages in the process are expected to manifest greater relative susceptibility in early life, whereas agents that act via later stage modes will tend to show greater susceptibility for exposures later in life.</p>
<p>Loeb LA & Harris CC. (2008). Advances in chemical carcinogenesis: a historical review and prospective. <i>Cancer Research</i> 68 (17): 6863-72.</p>	<p>This is an historical review and prospective on the advances in chemical carcinogenesis. The review starts from the first notes on chromatin changes by Theodor Boveri in 1914, and covers the step stones in molecular and cell biology research: the DNA discovery (1944), the DNA structure elucidation (1953), DNA sequencing and PCR method introduction (1977,1987) and many other key observations and experiments.</p>
<p>Hernández LG, van Steeg H, Luijten M & van Benthem J (2009). Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. <i>Mutation Research</i> 682: 94-109.</p>	<p>In order to better understand the mechanisms of known human non-genotoxic carcinogens and to illustrate the importance of a weight of evidence approach when evaluating their carcinogenic potential, the authors (1) evaluated the proportion of non-genotoxic carcinogens among known, probable and possible human carcinogens classified by the IARC, (2) estimated the risk of exposure of human non-genotoxic carcinogens through margin of exposure (MOE) evaluation, and (3) discussed potential alternative methods for their detection.</p> <p>The analysis demonstrated that human non-genotoxic carcinogens were present in 12% of IARC's Groups 1, 2A and 2B carcinogens and that a potential hazard was associated with 27% of them. The authors suggested that for all genotoxic chemicals, the mode of action is investigated for hazard and risk evaluation. Further, if negative genotoxic compounds have putative non-genotoxic modes of action, appropriate risk measures should be implemented.</p>
<p>López-Lázaro M (2010). A new view of carcinogenesis and an alternative approach to cancer therapy. <i>Molecular Medicine</i> 16 (3-4): 144-153.</p>	<p>This article reviews the evidence that tumour cells, in addition to acquiring a complex array of genetic changes, develop an alteration in the metabolism of oxygen. Although both changes play an essential role in carcinogenesis, the altered oxygen metabolism of cancer cells is not subject to the high genetic variability of tumours and may therefore be a more reliable target for cancer therapy. The author highlights the relevance of oxygen metabolism manipulation, as a novel approach, for the development of therapies that selectively target tumour cells.</p>
<p>Klaunig JE, Kamendulis LM & Hocevar BA (2010). Oxidative stress and oxidative damage in carcinogenesis. <i>Toxicologic Pathology</i> 38 (1): 96-109.</p>	<p>Chemical and physical agents including those that induce reactive oxygen species (ROS) can induce and/or modulate the multistep process of carcinogenesis. Based on the identification that oxidative damage to cellular macromolecules can arise through</p>

	<p>overproduction of ROS and faulty antioxidant and/or DNA repair mechanisms and ROS can stimulate signal transduction pathways and lead to activation of key transcription factors such as Nrf2 and NF-κB, the authors discuss the aspects of ROS biology in the context of their relationship to carcinogenesis. They also considered the recent evidence demonstrating an association between a number of single nucleotide polymorphisms (SNPs) in oxidative DNA repair genes and antioxidant genes with human cancer susceptibility.</p>
<p>Doull J, Cattley R, Elcombe C, Lake BG, Swenberg J, Wilkinson C, Williams G & van Gemert M (1999). A cancer risk assessment of di(2-ethylhexyl) phthalate: application of the new U.S. EPA risk assessment guidelines. <i>Regulatory Toxicology & Pharmacology</i> 29 (3): 327-357.</p>	<p>The authors used new toxicology data and a considerable amount of new mechanistic evidence to reconsider the cancer classification of DEHP under EPA's proposed new cancer risk assessment guidelines.</p> <p>The total weight-of-evidence clearly indicates that DEHP is not genotoxic. <i>In vivo</i> administration of DEHP to rats and mice results in peroxisome proliferation in the liver, and there is strong evidence and scientific consensus that, in rodents, peroxisome proliferation is directly associated with the onset of liver cancer. Recent studies of DEHP clearly indicate a nonlinear dose-response curve that strongly suggests the existence of a dose threshold below which tumours in rodents are not induced. Thus, the hepatocarcinogenic effects of DEHP in rodents result directly from the receptor-mediated, threshold-based mechanism of peroxisome proliferation, a well-understood process associated uniquely with rodents. Since humans are quite refractory to peroxisomal proliferation, even following exposure to potent proliferators such as hypolipidemic drugs, the authors concluded that the hepatocarcinogenic response of rodents to DEHP is not relevant to human cancer risk at any anticipated exposure level. DEHP should be classified an unlikely human carcinogen with a margin of exposure (MOE) approach to risk assessment.</p>
<p>Silva Lima B & Van der Laan JW (2000). Mechanisms of nongenotoxic carcinogenesis and assessment of the human hazard. <i>Regulatory Toxicology & Pharmacology</i> 32 (2): 135-43.</p>	<p>The purpose of this paper is to describe some mechanisms for nongenotoxic tumorigenicity and to indicate which type of testing should be done to substantiate why in those cases such a mechanism is not relevant to humans. The increasing attention being given to epigenetic carcinogenesis points at the need for a thorough evaluation during the toxicological program for safety assessment, enabling adequate assessment of the human hazard posed by such compounds. The authors suggest that data to support the nongenotoxic carcinogenesis may be obtained by collecting specific information from current safety assessment programs or from future, separate studies.</p>
<p>Zhang X & Ho SM (2011). Epigenetics meets endocrinology. <i>Journal of Molecular Endocrinology</i> February 1, 46 R11-R32.</p>	<p>A review on the relationship between epigenetic mechanisms and endocrine phenotypes. Specifically, the authors review the regulation by epigenetics of the three levels of hormone action (synthesis and release, circulating and target tissue levels, and target-organ responsiveness) and the epigenetic action of endocrine disruptors.</p>

<p>Labrecque MP, Takhar MK, Hollingshead BD, Prefontaine GG, Perdew GH & Beischlag TV (2012). Distinct roles for aryl hydrocarbon receptor nuclear translocator and ah receptor in estrogen-mediated signalling in human cancer cell lines. PLoS One 7 (1): e29545.</p>	<p>The activated AHR/ARNT complex (AHRC) regulates the expression of target genes upon exposure to environmental contaminants such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Importantly, evidence has shown that TCDD represses estrogen receptor (ER) target gene activation through the AHRC.</p> <p>The paper reports data indicating that AHR and ARNT act independently from each other at non-dioxin response element sites. The authors obtained experimental evidence demonstrating a dioxin-dependent repressor function for AHR and a dioxin-independent co-activator/co-repressor function for ARNT in estrogen signalling. The results provide with further insight into the mechanisms of transcription factor crosstalk and putative therapeutic targets in estrogen-positive cancers.</p>
<p>Lo PK & Sukumar S (2008). Epigenomics and breast cancer. Pharmacogenomics 9 (12): 1879-902.</p>	<p>The authors review known roles of the epigenetic machinery in the development and recurrence of breast cancer. Furthermore, they highlight the significance of epigenetic alterations as predictive biomarkers and as new targets of anticancer therapy.</p>
<p>Sadikovic B, Al-Romaih K, Squire JA & Zielenska M (2008). Cause and consequences of genetic and epigenetic alterations in human cancer. Current Genomics 9: 394-408.</p>	<p>While there has been considerable progress in understanding the impact of genetic and epigenetic mechanisms in tumourigenesis, there has been little consideration of the importance of the interplay between these two processes.</p> <p>In this review Sadikovic and co-workers summarized current understanding of the role of genetic and epigenetic alterations in human cancer. Furthermore, the authors considered the associated interactions of genetic and epigenetic processes in tumour onset and progression. They also provided a model of tumourigenesis to address the combined impact of both epigenetic and genetic alterations in cancer cells.</p>
<p>WHO, 2007 IPCS, Harmonization Project Document No. 4 Part 1: IPCS Framework for analysing the relevance of a cancer mode of action for humans and case-studies. http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf</p>	<p>This document presented a unified IPCS Human Relevance Framework for use of mode of action information in risk assessment for regulatory and other purposes, and it provided initial guidance for this task.</p>
<p>Non-standard methods and Integrated Approaches</p>	
<p>QSARs models for Carcinogenicity Prediction</p>	
<p>Woo YT, Lai DY, Argus MF & Arcos JC (1995). Development of structure–activity relationship rules for predicting carcinogenic potential of chemicals. Toxicology Letters 79: 219-228.</p>	<p>The authors describe the overall structure of the expert system, Oncologic, in the assessment of potential carcinogenic hazard of new chemicals discuss the scientific bases and principles of SAR analysis of chemical carcinogens used in the development of SAR knowledge rules, and delineate the major factors/rules useful for assessing the carcinogenic potential of fibers, polymers, metals/metalloids and several major classes of organic chemicals.</p> <p>An integrative approach using available short-term predictive tests and non-cancer toxicological</p>

	data to supplement SAR analysis has also been described.
Cronin MT, Jaworska JS, Walker JD, Comber MH, Watts CD & Worth AP. (2003). Use of QSARs in International Decision-Making Frameworks to Predict Health Effects of Chemical Substances Environmental Health Perspectives 111 (10): 1391-1401.	The authors reviewed the international regulatory use of QSARs to predict the health effects of chemical substances. Specifically, they described the recommendations of QSARs across several regulatory agencies to prioritize chemicals for testing and to fill data gaps in risk assessment data sets.
Woo YT & Lai DY (2003). Mechanisms of action of chemical carcinogens, and their role in Structure-Activity Relationships (SAR) analysis and risk assessment. In: Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens (R Benigni, ed). Boca Raton: CRC Press: 41-80.	This paper summarizes most of the available QSAR methods for the prediction of carcinogenicity.
Benigni R & Bossa C (2008). Predictivity of QSAR. Journal of Chemical Information & Modeling 48: 971-980.	The authors assessed the efficiency, potential and reliance of a range of QSAR methods for the prediction of mutagenicity and carcinogenicity.
Contrera JF, Kruhlak NL, Matthews EJ & Benz RD (2007). Comparison of MC4PC and MDL-QSAR rodent carcinogenicity predictions and the enhancement of predictive performance by combining QSAR models. Regulatory Toxicology & Pharmacology 49:172-182.	The aim of the study was to investigate on the interpretation and prediction of QSAR models by comparing the results of two software programs, MC4PC and MDL-QSAR, using the same 2-year rodent carcinogenicity training data set. The presented methodology was suggested to be particularly useful with endpoints such as rodent carcinogenicity.
Jaworska J & Nikolova-Jeliazkova N (2007). How can structural similarity analysis help in category formation? SAR and QSAR in Environmental Research 18: 195-207.	In this study the authors aimed to discuss the feasibility to apply computational structural similarity methods to augment formation of a category (filling safety data gaps).
Benigni R, Bossa C, Jeliazkova N, Netzeva N & Worth A (2008). The Benigni /Bossa rulebase for mutagenicity and carcinogenicity –a module of Toxtree. JRC Report, EUR 23241 EN – 2008.	This paper describes novel hazard estimation software called Toxtree, capable of making structure-based predictions for a number of toxicological endpoints, developed at the JRC. One of the modules developed as an extension to Toxtree is aimed at the prediction of carcinogenicity and mutagenicity. The main tool is a list of Structural Alerts (SA) for carcinogenicity.
Toropov AA, Toropova AP & Benfenati E (2009). Additive SMILES-based carcinogenicity models: probabilistic principles in the search for robust predictions. International	The present study was aimed to estimate the ability of the SMILES-based optimal descriptors to be a tool for QSAR analysis of carcinogenicity of non-congeneric chemicals.

Journal of Molecular Science 10: 3106-3127.	
Van Leeuwen K, Schultz TW, Henry T, Diderich B & Veith GD (2009). Using chemical categories to fill data gaps in hazard assessment. SAR QSAR Environmental Research 20: 207-220.	The authors of this paper, provided a summary of the growing use by regulatory agencies of the chemical categories approach, which groups chemicals based on their similar toxicological behaviour and fills in the data gaps in animal test data such as genotoxicity and aquatic toxicity.
Ferrari T & Gini G (2010). An open source multistep model to predict mutagenicity from statistical analysis and relevant structural alerts. Chemistry Central Journal 4 (Suppl 1), S2	The authors developed a cascade model validated on a large public set of molecular structures and their associated Salmonella mutagenicity outcome, for the prediction of mutagenicity. The data obtained approached the 85% reproducibility of the experimental mutagenicity Ames test and were reported on CEASAR website.
Fjodorova N, Vracko M, Novic M, Roncaglioni A & Benfenati E (2010). New public QSAR model for carcinogenicity. Chemistry Central Journal 4 Suppl 1:S3	The authors reviewed all the available QSAR models concluding that currently they cannot replace the human expert opinion and conventional methods. However, the combination of several methods is suggested to provide useful support to the overall evaluation of carcinogenicity. In the paper models for classification of carcinogenic compounds using MDL and Dragon descriptors were also developed.
Serafimova R, Fuart Gatnik M & Worth A (2010). Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. EUR 24427 EN. http://publications.jrc.ec.europa.eu/repository	The authors reviewed QSARs for genotoxicity and carcinogenicity, taking into accounts both models available in software tools and models that are published in the literature. They focused also on the potential applicability of diverse models to pesticides as well as to other types of regulated chemicals and pharmaceuticals.
Wu S, Blackburn K, Amburgey J, Jaworska J & Federle T (2010). A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regulatory Toxicology & Pharmacology 56: 67-81.	Wu et al. (2010) published a framework for using similarity based on chemical structure, reactivity, and metabolic and physicochemical properties to specifically evaluate the suitability of analogues for use in read-across toxicological assessments.
Threshold of Toxicological Concern	
U.S. FDA (1995) Department of Health and Human Services. Food Additives; Threshold of Regulation for Substances Used in Food-Contact Articles 21 CFR Parts 5, 25, 170, 171, and 174	Rules and Regulations on the consideration of a substance as a Food Additive and its use.

<p>Munro IC, Ford RA, Kennepohl E & Sprenger JG (1996). Correlation of structural class with no-observed effect levels: a proposal for establishing a threshold of concern. <i>Food & Chemical Toxicology</i> 34: 829-867.</p>	<p>Munro and co-workers explored the relationship between chemical structure and toxicity through the compilation of a large reference database consisting of over 600 chemical substances tested for a variety of endpoints. The study resulted in approximately 3000 no-observed-effect levels (NOELs). The authors built a database that was used to derive a threshold of acceptable human exposure for each of the structural classes that could be applied in the absence of specific toxicity data on a substance within one of the three structural classes.</p>
<p>Blackburn K, Stickney JA, Carlson-Lynch HL, McGinnis PA, Chappell L & Felter SP (2005). Application of the threshold of toxicological concern approach to ingredients in personal and household care products. <i>Regulatory Toxicology & Pharmacology</i> 43: 249-259.</p>	<p>The authors evaluated the applicability of the TTC database to ingredients used in consumer products based on a comparison of the diversity of chemical structures with those in the original TTC database and to confirm that the range of NOELs for these ingredients was consistent with the range of NOELs in the original database. The obtained results overlapped with the product ingredient structures and the NOELs for the ingredient chemicals were similar to the original dataset. Blackburn et al. supported the use of the TTC for ingredients in consumer products.</p>
<p>Kroes R, Renwick AG, Feron V, Galli CL, Gibney M, Greim H, Guy RH, Lhuguenot JC & Van de Sandt JJM (2007). Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. <i>Food & Chemical Toxicology</i> 45: 2533-2562.</p>	<p>This paper was aimed to apply the TTC methodology used for packaging migrants and flavouring agents to cosmetic ingredients and impurities. The analysis took into account the different types of exposure (oral vs. topical) with default conservative adjustment factors. The TTC approach was suggested to provide a useful additional tool for the safety evaluation of cosmetic ingredients and impurities of known chemical structure in the absence of chemical-specific toxicology data.</p>
<p>Safford RJ (2008). The dermal sensitization threshold - A TTC approach for allergic contact dermatitis. <i>Regulatory Toxicology & Pharmacology</i> 51: 195-200.</p>	<p>In this paper, Safford presents the outcomes of a probabilistic analysis of available sensitisation data, using data sets from ELINCS and a distribution for sensitisation potency from Local Lymph Node Assay data, in a similar way of TTCs. He established a Dermal Sensitisation Threshold (DST) was established below which it is suggested there is no appreciable risk of sensitisation, even for an untested ingredient. The author concludes that the use of DST would preclude the need for sensitisation testing of ingredients where dermal exposure is sufficiently low.</p>
<p>Carthew P, Clapp C & Gutsell S (2009). Exposure based waiving: the application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. <i>Food Chemical Toxicology</i> 47: 1287-1295.</p>	<p>In light of the concept of developing 'intelligent testing strategies' for REACH, the authors of this paper reviewed the inhalation toxicology studies available in the public domain to establish a database for inhalation toxicology and derive thresholds of toxicological concern (TTC) for effects in the respiratory tract. The authors propose that detailed exposure evaluation combined with potential QSAR for toxicity and TTCs for inhalation exposure could be used to waive [exposure based waiving (EBW) approach] undertaking inhalation toxicology studies under REACH.</p>
<p>Felter SP, Lane RW, Latulippe ME, Llewellyn GC, Olin SS, Scimeca JA & Trautman TD (2009). Refining the threshold of toxicological concern (TTC) for risk prioritization of trace</p>	<p>Felter and co-workers proposed the threshold of toxicological concern (TTC) methodology as a scientifically transparent approach for putting low-level exposures in the context of potential risk. Starting from the TTC literature the authors addressed to key two areas: genotoxicity data and time</p>

chemicals in food. Food and Chemical Toxicology 47: 2236-2245.	of exposure.
SCCP (2008), Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products SCCP/1171/08 (SCCS/SCHER/SCENIHR adopted on June 8, 2012).	The SCCP (now SCCS)/SCHER/SCENIHR (SCs) evaluated potential applications of the Threshold of Toxicological Concern (TTC) approach for human health risk assessment of chemical substances. The opinion focused on the potential applications of the TTC concept for cosmetics and other consumer products in relation to the mandates of the three SCs. The application of TTC in other areas, such as food, pharmaceuticals or EU chemical legislation (REACH) was not assessed, although such applications were described for completeness.
Bercu JP, Morton SM, Deahl JT, Gombar VK, Callis CM & van Lier RB (2010). In silico approaches to predicting cancer potency for risk assessment of genotoxic impurities in drug substances. Regulatory Toxicology & Pharmacology 57: 300-306.	The authors proposed the use of TTC in combination with QSAR tools to establish safe levels for genotoxic impurities (GTIs) in drug substances. Indeed, the aim of the study was to use in silico tools to predict the cancer potency (TD(50)) of a compound based on its structure. Structure activity relationship (SAR) models were developed from the carcinogenicity potency database using MultiCASE and VISDOM softwares. The approach for non-potent compounds was considered useful to establish safe levels greater than the TTC for GTIs in a drug substance.
LED for <i>in vivo</i> Genotoxicity to estimate Cancer Potency	
Sanner T & Dybing E (2005). Comparison of carcinogenic and <i>in vivo</i> genotoxic potency estimates. Basic Clinical Pharmacology & Toxicology 96: 131-139.	The finding of a linear relationship between the lowest effective dose (LED) for <i>in vivo</i> genotoxicity and the carcinogen dose descriptor T25 is of importance. It was found for 34 carcinogens studied which covered a potency range of 10,000, that the median of the ratio LED/T25 was equal to 1.05 and that for 90% of the substances the numerical value of LED was similar to the numerical value of T25 within a factor of less than 5–10. The results suggest that if further evaluated, LED for <i>in vivo</i> genotoxicity may be used in a semi-quantitative method for risk assessment of mutagens without a long-term study.
Hernández LG, Slob W, van Steeg H & van Benthem J (2011). Can carcinogenic potency be predicted from <i>in vivo</i> genotoxicity data?: a meta-analysis of historical data. Environmental & Molecular Mutagenesis 52 (7): 518-28.	A preliminary investigation on the applicability of <i>in vivo</i> genotoxicity tests to estimate cancer potency looked promising. Positive correlations between <i>in vivo</i> genotoxicity (micronucleus test and transgenic rodent mutation test) and carcinogenic potency were found. Dose-response analyses using sophisticated dose-response software such as PROAST (RIVM) or the BMDS (USEPA) were used. The results suggest that <i>in vivo</i> genotoxicity tests may be used to estimate carcinogenic potency.
Hernández LG, van Benthem J & Johnson GE (2013). A mode-of-action approach for the identification of genotoxic carcinogens. PLoS One May 13; 8 (5): e64532.	In this study the authors performed the cytokinesis-block micronucleus assay and immunofluorescence mitotic machinery visualisation in human lymphoblastoid (AHH-1) and

	<p>Chinese Hamster fibroblast (V79) cell lines after treatment with the aneugen 17-β-oestradiol (E₂).</p> <p>Results were compared to previously published data on bisphenol-A (BPA) and Rotenone data. Two concentration-response approaches (the threshold-[Td] and benchmark-dose [BMD] approaches) were applied to derive a point of departure (POD) for in vitro MN induction. Ranking comparisons of the PODs from the in vitro MN and the carcinogenicity studies demonstrated a link between these two endpoints for BPA, E₂ and Rotenone. The analysis was extended to include additional aneugens, clastogens and mutagens and further concentration and dose-response correlations were observed between PODs from the in vitro MN and carcinogenicity. The approach looked promising and the authors suggested it could be extended to other genotoxic carcinogens. The MOA and quantitative information from the in vitro MN studies could be used in a quantitative manner to further inform cancer risk assessment.</p>
Cell Transformation Assay (CTA)	
<p>DiPaolo JA, Donovan P & Nelson R (1969). Quantitative studies of <i>in vitro</i> transformation by chemical carcinogens. Journal of the National Cancer Institute 42: 867-874.</p>	<p>The authors report that chemical carcinogens can induce malignant transformation in mammalian cell systems. However, the precise molecular and cellular alterations that result in transformation have not been identified.</p>
<p>Isfort RJ, Kerckaert GA & LeBoeuf RA (1996). Comparison of the standard and reduced pH Syrian hamster embryo (SHE) cell <i>in vitro</i> transformation assays in predicting the carcinogenic potential of chemicals. Mutation Research 356: 11-63.</p>	<p>The article is a comprehensive review of the Syrian Hamster Embryo (SHE) cell transformation literature; it was performed in order to catalogue the chemical/physical entities which have been evaluated for <i>in vitro</i> cell transformation potential. Both reduced pH (pH 6.7) and standard pH (pH 7.1-7.3) SHE cell testing protocols were considered. The authors suggest that the SHE cell transformation assay is predictive for rodent carcinogenicity under either reduced or standard pH conditions; with the assay displaying better performance in prediction capability under reduced pH conditions.</p>
<p>Matthews EJ, Spalding JW & Tennant RW (1993). Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays.</p>	<p>This report describes the activities of 168 chemicals tested in a standard transformation assay using A-31-1-13 BALB/c-3T3 cells.</p> <p>Data analyses revealed that the transformation assay and rodent bioassay had a concordance of 71%; sensitivity for carcinogens of 80.0%; specificity for detecting non carcinogens of 60%.</p>

<p>Environmental Health Perspectives 101 Suppl 2: 347-482.</p>	<p>In contrast, Salmonella mutagenicity assays and rodent bioassays had a concordance of 63%, a sensitivity of 58%, and a specificity of 69%. The transformation assay complemented the Salmonella mutagenesis assay in the identification of nonmutagenic carcinogens; thus, the two assays had a combined 83% sensitivity for all carcinogens and a specificity of 75% for non-mutagenic non carcinogens.</p>
<p>LeBoeuf RA, Kerckaert KA, Aardema MJ & Isfort RJ (1999). Use of Syrian hamster embryo and BALB/c 3T3 cell transformation for assessing the carcinogenic potential of chemicals. IARC Science Publications 146: 409-25.</p>	<p>CTAs have been shown to involve a multistage process that closely models key stages of carcinogenesis.</p>
<p>Jacobson-Kram D & Jacobs A (2005). Use of genotoxicity data to support clinical trials or positive genotox findings on a candidate pharmaceutical or impurity ... now what? International Journal of Toxicology. 24 (3): 129-34.</p>	<p>In situations where a genetic toxicology assay showed a positive result, some review divisions have asked sponsors to perform a Syrian hamster embryo (SHE) cell transformation assay or a p53 carcinogenicity study prior to allowing repeat-dose clinical trials to proceed. This paper discusses alternatives to SHE cell and p53 assays when faced with a positive result in a genetic toxicology assay. In addition, the authors discuss factors to consider when setting limits for genotoxic impurities in drug substances and products.</p>
<p>Poth A, Heppenheimer A & Bohnenberger S. (2007). Bhas42 cell transformation assay as a predictor of carcinogenicity. ALTEX. 14: 519-521.</p>	<p>Biomarkers associated with human cancer have been studied as a means of obtaining a more objective measure of cell transformation in the Bhas 42 assay. The authors observed in transformed Bhas 42 cells a significant increase of expression and activity of cholinesterase and alkaline phosphatase, respectively, compared to non-transformed cells. Very high activity and intensity were found in Type III foci compared with Type II.</p>
<p>Ao L, Liu JY, Liu WB, Gao LH, Hu R, Fang ZJ, Zhen ZX, Huang MH, Yang MS & Cao J (2010). Comparison of gene expression profiles in BALB/c 3T3 transformed foci exposed to tumor promoting agents. Toxicology <i>in Vitro</i> 24: 430-438.</p>	<p>The authors of this study used cDNA microarrays to analyse gene expression profiles and discern chemical-associated profiles induced by a variety of tumour promoting agents in transformed BALB/c 3T3 cells. The cDNA microarray was based on 1796 mouse genes. The expression of genes associated to specific pathways was detected. Indeed, genes involved in oxidative stress response or cell proliferation were up-regulated. The application of gene expression profile analysis was suggested to be extremely useful in providing insights into the etiology of different chemicals.</p>
<p>Pfuhler S, Kirst A, Aardema M, Banduhn N, Goebel C, Araki D, Costabel-Farkas M, Dufour E, Fautz R, Harvey J, Hewitt NJ, Hibatallah J, Carmichael P, Macfarlane M, Reisinger K, Rowland J, Schellauf F, Schepky A & Scheel J (2010). A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: Genotoxicity. A COLIPA</p>	<p>This article described common approaches of cosmetic companies, with recommendations for evaluating <i>in vitro</i> genotoxins using non-animal approaches. A weight of evidence approach was employed to set up a decision-tree for the integration of alternative methods into tiered testing strategies.</p>

analysis. <i>Regulatory Toxicology & Pharmacology</i> 57: 315-324.	
Vanparys Ph, Corvi R, Aardema M, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2010). ECVAM prevalidation of three cell transformation assays. <i>ALTEX</i> 27: 267-270.	A prevalidation study on the cell transformation assays in SHE cells at pH 6.7, SHE cells at pH 7.0 and Balb/c 3T3 cell line was coordinated by ECVAM focussing on issues of standardisation of protocols, within-laboratory reproducibility, test method transferability and between-laboratory reproducibility. The Validation Management Team concluded that standardised protocols are now available that should be the basis for future use. The SHE pH 6.7, and the SHE pH 7.0 protocols and the assays system themselves are transferable between laboratories, and are reproducible within- and between-laboratories. For the Balb/c 3T3 method, some clarifications and modifications to the protocol were needed to obtain reproducible results. Overall, three methods have shown to be valuable to detect rodent carcinogens.
Sakai A, Sasaki K, Hayashi K, Muramatsu D, Arai S, Endou N, Kuroda S, Albrecht P, Bohnenberger S, Kunkelmann T, Asakura M, Hirose H, Ishii N, Mizuhashi F, Kasamoto S, Nagai M, Pant K, Bruce SW, Sly JE, Yamazaki S, Umeda M & Tanaka N (2011). An international validation study of a Bhas 42 cell transformation assay for the prediction of chemical carcinogenicity. <i>Mutation Research</i> 725: 57-77.	An international study was carried out to validate the Bhas 42 cell transformation assay in which six laboratories from three countries participated. The Bhas 42 cell transformation assay is a sensitive short-term system for predicting chemical carcinogenicity. Bhas 42 cells were established from BALB/c 3T3 cells by the transfection of v-Ha-ras gene and postulated to have acquired an initiated state in the two-stage carcinogenesis theory. The Bhas 42 cell transformation assay is capable of detecting both tumour-initiating and tumour-promoting activities of chemical carcinogens. Here, the results obtained are reported. The authors demonstrated that the Bhas 42 cell transformation assay is transferable and reproducible between laboratories and applicable to the prediction of chemical carcinogenicity.
Jacquet N, Maire MA, Landkocz Y, & Vasseur P (2012). Carcinogenic potency of perfluorooctane sulfonate (PFOS) on Syrian hamster embryo (SHE) cells. <i>Archives of Toxicology</i> 86 (2): 305-14.	The authors assessed carcinogenic potential of Perfluorooctane sulfonate (PFOS) by studying morphological transformation in Syrian hamster embryo (SHE) cells. Genotoxicity of PFOS and expression of PPARs genes in SHE cells were also measured. The results indicated that PFOS behave as a non-genotoxic carcinogen and impacted PPARs genes. Interestingly, its cell transforming potential paralleled an increased expression of ppar- β/δ .
Vanparys P, Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S & Schechtman L (2012). Application of <i>in vitro</i> cell transformation assays in regulatory toxicology for pharmaceuticals, chemical, food products and cosmetics. <i>Mutation Research</i> 744 (1): 111-116.	The extensive review on CTAs by the OECD (2007) and the proven standardisation, intra- and inter-laboratory reproducibility of the SHE CTAs justifies broader use of these methods to assess carcinogenic potential of chemicals. The manuscript describes possible applications of the CTA in relation to different industrial sectors, e.g. food additives, chemicals, cosmetics, pharmaceuticals, etc.
Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S, Schechtman L & Vanparys P. (2012). ECVAM prevalidation	A prevalidation study of the <i>in vitro</i> cell transformation assays (CTAs) using the BALB/c 3T3 cell line, SHE cells at pH 6.7, and SHE cells at pH 7.0 was coordinated by the European Centre for the

<p>study on <i>in vitro</i> cell transformation assays: general outline and conclusions of the study. Mutation Research 744 (1): 12-9.</p>	<p>Validation of Alternative Methods (ECVAM) and focused on issues of standardisation of protocols, test method transferability and within- and between-laboratory reproducibility. The paper reports the results of the ECVAM study demonstrating that for the BALB/c 3T3 method, some modifications to the protocol were needed to obtain reproducible results between laboratories, while the SHE pH 6.7 and the SHE pH 7.0 protocols are transferable between laboratories, and results are reproducible within- and between-laboratories.</p> <p>The authors recommend that the BALB/c 3T3 and SHE protocols as instituted in this prevalidation study should be used in future applications of these respective transformation assays. They also recommend the development of an OECD test guideline for the SHE CTAs, based on the protocol published in this issue to support the harmonised use and regulatory application of the methods. The development of an OECD test guideline for the BALB/c 3T3 CTA should likewise be further pursued upon the availability of additional supportive data and improvement of the statistical analysis.</p>
<p>Benigni R, Bossa C & Tcheremenskaia O (2013). In vitro cell transformation assays for an integrated, alternative assessment of carcinogenicity: a data-based analysis. Mutagenesis 28 (1): 107-16.</p>	<p>Based on the evidence that the 2-year bioassay can fail to detect all genotoxic carcinogens and it cannot detect nongenotoxic carcinogens, the authors report on an integrated strategy consisting of the <i>in vitro</i> Ames and SHE cells transformation assays, combined with QSARs, as a valid alternative to the present pre-screening strategies. The analysis has furthered by (i) including results of CTAs on inorganics, together with and the Bhas42 assay and (ii) considering new structural alerts for nongenotoxic carcinogenicity. The author suggest the strategy as an efficient tool to identify both genotoxic and nongenotoxic carcinogens.</p>
<p>OECD, (2007) no. 31 Detailed Review Paper on cell transformation assays for detection of chemical carcinogens, no. 31 ENV/JM/MONO (2007)18.</p>	<p>This document focuses on the three main <i>in vitro</i> CTAs, the SHE, the BALB/c 3T3 and the C3H10T1/2 assays. The SHE assay uses karyotypically normal cells and is believed to detect early steps of carcinogenesis. The other two assays are based on immortalized aneuploid cell lines which measure later stages of carcinogenesis. The DRP is a comprehensive collection of available data reported in the literature for three CTAs which are reported alongside genotoxicity test data using mammalian and non-mammalian cell systems. The performances of the CTAs in predicting carcinogenic potential are analysed in terms of classification as rodent and/or human carcinogens. Based on their performances, the OECD recommended that the CTAs using the SHE cells and the BALB/c 3T3 cell line be developed into OECD Test Guidelines.</p>
<p>EURL ECVAM (2012) Recommendation concerning the cell transformation assays using Syrian hamster embryo cells (SHE) and the BALB/c 3T3 mouse fibroblast cell line for <i>in vitro</i> carcinogenicity testing. Annex I: ESAC opinion on the</p>	<p>EURL ECVAM recommendation report on the three CTA protocol variants: (a) The SHE CTA performed at pH 6.7 (SHE pH 6.7 CTA); (b) the SHE CTA performed at pH 7.0 (SHE pH 7.0 CTA); and (c) the BALB/c 3T3 CTA. ECVAM coordinated an international study that was designed to address issues of CTA protocols</p>

<p>ESAC peer review of an ECVAM-coordinated prevalidation study concerning three protocols of the cell transformation assay (CTA) for <i>in vitro</i> carcinogenicity testing. http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing</p>	<p>standardisation, transferability and reproducibility in three protocol variants for the SHE CTA (at pH 6.7 and pH 7.0) and the BALB/c 3T3 assay. The study of the three test methods was peer reviewed by the ESAC that issued a peer review report and an ESAC opinion. Based on the available documents (study reports, ESAC opinion, OECD DRP), EURL ECVAM recommended that an OECD TG be developed for the SHE CTA and that further investigations were needed to confirm the performance of the BALB/c 3T3 CTA.</p> <p>Moreover, the recommendation stated that in conjunction with other available data, the CTAs have the potential of partial replacement or reduction when used in a weight of evidence approach for hazard identification and risk assessment.</p>
<p>EFSA (2011). Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011; 9 (9): 2379 (69pp)</p>	<p>The Scientific Committee reviewed the current state-of-the-science on genotoxicity testing and provided a commentary and recommendations on genotoxicity testing strategies. A step-wise approach is recommended for the generation and evaluation of data on genotoxic potential, beginning with a basic battery of <i>in vitro</i> tests, comprising a bacterial reverse mutation assay and an <i>in vitro</i> micronucleus assay. It is suggested to consider whether specific features of the test substance might require substitution of one or more of the recommended <i>in vitro</i> tests by other <i>in vitro</i> or <i>in vivo</i> tests in the basic battery. In the event of negative <i>in vitro</i> results, it can be concluded that the substance has no genotoxic potential. In case of inconclusive, contradictory or equivocal results, it may be appropriate to conduct further testing <i>in vitro</i>. In case of positive <i>in vitro</i> results, review of the available relevant data on the test substance and, where necessary, an appropriate <i>in vivo</i> study to assess whether the genotoxic potential observed <i>in vitro</i> is expressed <i>in vivo</i> is recommended. The approach to <i>in vivo</i> testing should be also step-wise. The combination of assessing different endpoints in different tissues in the same animal <i>in vivo</i> should also be considered.</p>
<p>Oxidative Stress and GJIC Communication Measurements</p>	
<p>Klaunig JE & Shi Y (2009). Assessment of gap junctional intercellular communication. Current Protocols in Toxicology. Chapter 2:Unit 2.17.</p>	<p>Taking into account the key role of gap-junctions aberrations in the development of several human diseases such as cancer, cardiac arrhythmia, Charcot-Marie-tooth disease, and viscerotaxial heterotaxia syndrome, Klaunig and Shi describe methods for measuring gap junctional intercellular communication using primary mouse hepatocytes as a model. The authors focus on functional evaluation based on dye coupling.</p>
<p>Klaunig JE, Kamendulis LM & Hocevar BA (2010). Oxidative stress and oxidative damage in carcinogenesis. Toxicologic Pathology 38 (1): 96-109.</p>	<p>Chemical and physical agents including those that induce reactive oxygen species (ROS) can induce and/or modulate the multistep process of carcinogenesis. Based on the identification that oxidative damage to cellular macromolecules can arise through overproduction of ROS and faulty antioxidant and/or DNA repair mechanisms and ROS can</p>

	stimulate signal transduction pathways and lead to activation of key transcription factors such as Nrf2 and NF-kappaB, the authors discuss the aspects of ROS biology in the context of their relationship to carcinogenesis. They also considered the recent evidence demonstrating an association between a number of single nucleotide polymorphisms (SNPs) in oxidative DNA repair genes and antioxidant genes with human cancer susceptibility.
Toxicogenomics-based tests and <i>The carcinoGENOMICS Project</i>	
Tsujimura K, Asamoto M, Suzuki S, Hokaiwado N, Ogawa K & Shirai T (2006). Prediction of carcinogenic potential by a toxicogenomic approach using rat hepatoma cells. <i>Cancer Science</i> 97: 1002-1010.	The authors conducted gene expression profiling of cultured rat hepatoma cells (MH1C1) exposed to carcinogenic chemicals with the aim of providing a basis for rapid and reliable prediction of carcinogenicity using microarray technology. Cells were treated with chemicals for three days at a non-toxic dose and the analysis of gene expression profile with an in-house microarray allowed a set of genes to be identified differentiating hepatocarcinogens from non-carcinogens, and all carcinogens from non-carcinogens, by statistical methods. The authors sought promising outcomes from these short-term bioassay systems for carcinogenicity.
Guo L, Lobenhofer EK, Wang C, Shippy R, Harris SC, Zhang L, Mei N, Chen T, Herman D, Goodsaid FM, Hurban P, Philips KL, Qu J, Deng X, Sun YA, Tong W, Dragon WT & Shi L (2006). Rat toxicogenomic study reveals analytical consistency across microarray platforms. <i>Nature Biotechnology</i> 24: 1162-1169.	In the context of the MicroArray Quality Control (MAQC) project, the authors validated and extended the obtained data. Toxicogenomics data sets were built using 36 RNA samples from rats treated with three chemicals (aristolochic acid, riddelliine and comfrey) and each sample was hybridized to four microarray platforms. The real-world toxicogenomic data set reported here showed high concordance in intersite and cross-platform comparisons.
Nie AY, McMillian M, Parker JB, Leone A, Bryant S, Yieh L, Bittner A, Nelson J, Carmen A, Wan J & Lord PG (2006). Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. <i>Molecular Carcinogenesis</i> 45: 914-933.	Nie and co-workers built a large gene expression database using cDNA microarrays and liver samples treated with over one hundred paradigm compounds; data were obtained from male rats treated for 24 hrs. The aim of the study was to determine gene expression signatures for nongenotoxic carcinogens (NGTCs). The obtained gene expression profiles accurately predicted NGTC potential of compounds and resulted amenable for other toxicity signatures.
Li HH, Aubrecht J & Fornace AJ Jr (2007). Toxicogenomics: overview and potential applications for the study of non-covalent DNA interacting chemicals. <i>Mutation Research</i> 623: 98-108.	In this review the authors focus their attention onto the characterization of stress agents including non-covalent DNA interacting chemicals based on gene expression profiling, through the use of functional genomics approaches, highlighting both pro and cons of these methodologies.
Le Fevre AC, Boitier E, Marchandau JP, Sarasin A & Thybaud V (2007). Characterization of DNA reactive and	In this study, human lymphoblastoid TK6 cells were exposed to several anticancer drugs for 4-h and examined them immediately or after a 20-h recovery period. Cytotoxicity (cell cycle), genotoxicity

<p>non-DNA reactive anticancer drugs by gene expression profiling. <i>Mutation Research</i> 619: 16-29.</p>	<p>(micronucleus), and gene expression profiles (human Affymetrix U133A GeneChips-22K) were evaluated. 28 marker genes mainly involved in signal transduction and cell cycle pathways were detected, which were suggested to be useful as a predictive model to classify genotoxins according to their direct or indirect interaction with DNA.</p>
<p>Stemmer K, Ellinger Ziegelbauer H, Ahr HJ & Dietrich DR (2007). Carcinogen-specific gene expression profiles in short-term treated Eker and wild-type rats indicative of pathways involved in renal tumorigenesis. <i>Cancer Research</i> 67: 4052-4068.</p>	<p>In this study, Eker rats, heterozygous for a dominant germ line mutation in the tuberous sclerosis 2 (Tsc2) tumour suppressor gene, were used as a model to study renal carcinogenesis. Eker and corresponding wild-type rats were exposed to genotoxic aristolochic acid (AA) or non-genotoxic ochratoxin A (OTA) to elucidate early carcinogen-specific gene expression changes and to test whether Eker rats were more sensitive to carcinogen-induced changes in gene expression.</p>
<p>Fielden MR, Brennan R & Gollub J (2007). A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. <i>Toxicological Sciences</i> 99: 90-10.</p>	<p>Fielden and co-workers developed a multigene biomarker (signature) to predict the likelihood of non-genotoxic chemicals to induce liver tumours in long-term studies. The authors used hepatic gene expression profiles from rats 5-days-treated with several non-genotoxic hepatocarcinogens and non hepatocarcinogens (100). The authors found this method very accurate and able to facilitate the identification of mode of actions. Indeed, the signature was suggested to be helpful for the early assessment of cancer risk.</p>
<p>Nioi P, Pardo ID, Sherratt PJ & Snyder RD (2008). Prediction of non-genotoxic carcinogenesis in rats using changes in gene expression following acute dosing. <i>Chemico-Biological Interactions</i> 172: 206-215.</p>	<p>In this study the authors investigated whether early changes in gene expression may be developed as markers for cancer. Gene expression profiles, from animals treated with either non-genotoxic carcinogens or non-carcinogens, were analysed by qPCR. The data obtained suggested that acute changes in gene expression might be helpful to identify biomarkers for non-genotoxic carcinogenesis.</p>
<p>Uehara T, Hirode M, Ono A, Kiyosawa N, Omura K, Shimizu T, Mizukawa Y, Miyagishima T, Nagao T & Urushidani T (2008). A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. <i>Toxicology</i> 250: 15-26.</p>	<p>In this study, the possibility of applying a toxicogenomics approach by using comprehensive gene expression data in rat liver treated with various compounds was explored.</p>
<p>Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A & Ahr HJ (2008). Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term <i>in vivo</i> studies. <i>Mutation Research</i> 637: 23-39.</p>	<p>The authors evaluated the carcinogenic potential of several different chemicals (genotoxic and non-genotoxic carcinogens and hepatocarcinogens and from non-carcinogens) using gene expression profile from livers of 14-days-treated rats. The data obtained were considered accurate enough to proof the feasibility of this method as classifier of carcinogens in short-term studies.</p>
<p>Guyton KZ, Kyle AD, Aubrecht J, Cogliano VJ, Eastmond</p>	<p>In light of the difficulties at discriminating between genotoxic and non-genotoxic carcinogens and</p>

<p>DA, Jackson M, Keshava N, Sandy MS, Sonawane B, Zhang L, Waters MD & Smith MT (2009). Improving prediction of chemical carcinogenicity by considering multiple mechanisms and applying toxicogenomic approaches. <i>Mutation Research</i> 681: 230-240.</p>	<p>the at identifying the cancer dose-response relationship at low doses, Guyton et al report on new bioanalytical approaches (e.g., transcriptomics, proteomics, and metabolomics) which could be applied in human, animal and <i>in vitro</i> studies and could better characterize a wider array of hazard traits. Thus, to improving the ability to predict the potential carcinogenicity of chemicals.</p>
<p>Ellinger-Ziegelbauer H, Aubrecht J, Kleinjans JC & Ahr HJ (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. <i>Toxicology Letters</i> 186: 36-44.</p>	<p>The paper underlies the importance of developing novel scientific approaches bridging genotoxicity and carcinogenicity testing via understanding underlying mechanisms for facilitating cancer risk assessment. The authors consider extremely promising the toxicogenomics approaches as these have the potential of providing generic insight in molecular pathway responses. The goal of the report is to review recent progress in the development and application of toxicogenomics to the derivation of genomic biomarkers associated with mechanisms of genotoxicity and carcinogenesis.</p>
<p>Jonker MJ, Bruning O, van Iterson M, Schaap MM, van der Hoeven TV, Vrieling H, Beems RB, de Vries A, van Steeg H, Breit TM & Luijten M (2009). Finding transcriptomics biomarkers for <i>in vivo</i> identification of (non-)genotoxic carcinogens using wild-type and Xpa/p53 mutant mouse models. <i>Carcinogenesis</i> 30: 1805-1812.</p>	<p>In this study the authors treated both wild-type and DNA repair-deficient Xpa(-/-)/p53(+/-) (Xpa/p53) mice up to 14 days to compounds of three distinct classes: genotoxic carcinogens (GTXC), non-genotoxic carcinogens (NGTXC) and non-carcinogens. Subsequently, extensive transcriptomics analyses were performed on several tissues, and transcriptomics data were screened for potential biomarkers using advanced statistical learning techniques. Jonker and collaborators obtained highly confident results and demonstrated a proof-of-principle for the identification and use of transcriptomics biomarkers for GTXC or NGTXC, for short-term studies.</p>
<p>Mathijs K, Brauers KJ, Jennen DGJ, Lizarraga D, Kleinjans JCS & van Delft JHM (2010). Gene expression profiling in primary mouse hepatocytes discriminates true from false-positive genotoxic compounds. <i>Mutagenesis</i> 25: 561-568.</p>	<p>Considering the high rate of false-positive for predicting <i>in vivo</i> genotoxicity and carcinogenicity obtained from standard <i>in vitro</i> tests (Ames, CAvit MNvit), Mathijs and collaborators investigated the potential of gene expression profiling in metabolically competent primary mouse hepatocytes to discriminating true genotoxic (GTX) compounds from false-positive genotoxic (FP-GTX) compounds.</p>
<p>Doktorova TY, Pauwels M, Vinken M, Vanhaecke T & Rogiers V (2012) Opportunities for an alternative integrating testing strategy for carcinogen hazard assessment? <i>Critical Reviews in Toxicology</i> 42 (2): 91-106.</p>	<p>In this review paper, the major advantages and pitfalls of the existing alternative methodologies to the carcinogenicity bioassay are discussed. Based on the available scientific data in the public domain, the authors propose a "feasible integrated testing strategy" which incorporates some promising alternatives (toxicogenomics-, transcriptomics-based tests and CTAs), providing at the same time information on the mechanism of action and the toxic nature of the compounds tested.</p>
<p>Doktorova TY, Yildirimman R, Vinken M, Vilardell M, Vanhaecke T, Gmuender H, Bort R, Brolen G, Holmgren G, Li R, Chesne C, van Delft J, Kleinjans J, Castell J, BJORQUIST</p>	<p>In the present study, the transcriptomics responses following exposure to genotoxic and non-genotoxic hepatocarcinogens and non-carcinogens in five liver-based <i>in vitro</i> models, namely conventional and epigenetically stabilized cultures of primary rat hepatocytes, the human hepatoma-</p>

<p>P, Herwig R, & Rogiers V (2013). Transcriptomic responses generated by hepatocarcinogens in a battery of liver-based <i>in vitro</i> models. <i>Carcinogenesis</i> 34: 1393-1402.</p>	<p>derived cell lines HepaRG and HepG2 and human embryonic stem cell-derived hepatocyte-like cells, are examined.</p>
<p>Jennings P, Limonciel A, Felice L & Leonard MO (2013). An overview of transcriptional regulation in response to toxicological insult. <i>Archives of Toxicology</i> 87: 49–72.</p>	<p>This review aims to consolidate and summarise the major toxicologically relevant transcription factor-governed molecular pathways. The authors focus on the importance of toxicogenomic data sets and information on regulation of stress pathways at the transcriptome level, in order to appreciate the diversity and complexity of biological responses to xenobiotics.</p>
<p>Waters MD, Jackson M & Lea I (2010). Characterizing and Predicting Carcinogenicity and Mode of Action Using Conventional and Toxicogenomics Methods. <i>Mutation Research</i> 705: 184-200.</p>	<p>The authors report that the evidence accumulated to date suggests that gene expression profiles reflect underlying modes or mechanisms of action, such that they will be useful in the prediction of chemical carcinogenicity, especially in conjunction with conventional short-term tests for gene mutation, chromosomal aberration and aneuploidy.</p>
<p>Rohrbeck A, Salinas G, Maaser K, Linge J, Salovaara S, Corvi R & Borlak J (2010). Toxicogenomics applied to <i>in vitro</i> carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens. <i>Toxicological Sciences</i> 118: 31-41.</p>	<p>The authors reported the application of advanced genomics to a cellular transformation assay to identify toxicity pathways and gene signatures predictive for carcinogenicity. During the study, the genome-wide gene expression analysis and quantitative real time polymerase chain reaction (qRT-PCR) were applied to untransformed and transformed mouse fibroblast Balb/c 3T3 cells that were exposed to either 2, 4-diaminotoluene, benzo(a)pyrene, 2-acetylaminoflourene, or 3-methycolanthrene at IC20 conditions for 24 and 120 h, respectively. Prediction of carcinogenicity potential of the three substances was significant and the testing strategy was able to identify commonly regulated carcinogenic pathways and a gene signature that predicted the risk for carcinogenicity.</p>
<p>Vinken M, Vanhaecke T & Rogiers V (2012). Primary hepatocyte cultures as <i>in vitro</i> tools for toxicity testing: quo vadis? <i>Toxicology in Vitro</i> 26: 541-544.</p>	<p>The authors discuss on the mechanisms underlying hepatocyte dedifferentiation, which in turn can affect primary hepatocytes cultivation for <i>in vitro</i> toxicity testing. Recent insights into this mechanism have been suggested to be useful for the development of new long-term hepatic <i>in vitro</i> methods, especially in the light of the stringent European legislative modifications.</p>
<p>Transgenic mouse models</p>	
<p>Tennant RW, French JE & Spalding JW (1995). Identifying chemical carcinogens and assessing potential risk in short-term bioassays using transgenic mouse models. <i>Environmental Health Perspectives</i> 103: 942-950.</p>	<p>Tenant et al reported on the advantages of transgenic mouse models for the carcinogenicity testing. In particular, the authors showed that mutagenic carcinogens could be identified with increased sensitivity and specificity using hemizygous p53 mice in which one allele of the p53 gene has been inactivated. The TG.AC transgenic model, carrying a v-Ha-ras construct, responded to a number of mutagenic and nonmutagenic carcinogens and tumour promoters, but not to non carcinogens, as well.</p>

	A decision-tree approach for testing chemicals was also presented.
Tennant RW, Stasiewicz S, Mennear J, French JE & Spalding JW (1999). Genetically altered mouse models for identifying carcinogens. IARC Scientific Publications 146: 123-150.	Modifications to the mouse genome can induce increased sensitivity to tumour incidence and formation. The possibility to identify carcinogens through transgenic mouse models relies on these modifications.
Eastin WC, Haseman JK, Mahler JF & Bucher JR (1998). The National Toxicology program evaluation of genetically altered mice as predictive models for identifying carcinogens. Toxicologic Pathology 26: 461-473.	The authors report on the evaluation, performed by the National Toxicology Program, of two mouse models potentialities, the Tg.AC (carrier of an activated mouse H-ras oncogene) and the p53+/- (heterozygous for the wild-type tumor suppressor gene Trp53) as a substitute for the 2-year rodent assays.
Bucher JR (1998). Update on National Toxicology Program (NTP) assays with genetically altered or “transgenic” mice. Environmental Health Perspectives 106: 619-621.	Brief review on the advance of transgenic mouse models research and development, based on the evaluation of several lines of genetically modified mice, performed by the National Toxicology Program.
Ashby J (2001). Expectations for transgenic rodent cancer bioassay models. Toxicologic Pathology 29 Suppl, 177-182.	In this paper, Ashby discussed about his analysis on the transgenic mouse models: advances in development, validation and potential to standard rodent bioassay replacement. The author focused on the phases in need of being improved for a more critical evaluation of new models. It was suggested further discussion of 5 key areas from test methodology to chemicals databases analysis, carcinogenesis process study to regulations and legislations update.
Marx J (2003). Building Better Mouse Models for Studying Cancer. Science 299: 1972-1975.	The article highlights the advances on new strains of transgenic mice research: in particular, mice baring tumours that more closely mimic human cancers, for use in both basic cancer studies and drug screening.
Pritchard JB, French JE, Davis BJ & Haseman JK (2003). The role of transgenic mouse models in carcinogen identification. Environmental Health Perspectives 111: 444-454.	The paper examined existing data on the use of transgenic mouse models for identification of human carcinogens. The authors on the three most extensively studied of these mice, Trp53+/-, Tg/AC, and RasH2, and compare their performance with the traditional 2-year rodent bioassay. Data on 99 chemicals were evaluated. A variety of potential testing strategies ranging from individual transgenic models to combinations of these three models with each other and with traditional rodent assays was evaluated using IARC carcinogens database. Transgenic models, for their performance, were suggested for regulatory acceptance and use in human cancer risk assessment, after further validation and standardization.
De Vries A, van Steeg H & Opperhuizen A (2004). Transgenic mice as alternatives in carcinogenicity testing: current status. RIVM report 340700001	Transgenic mouse models showed a high specificity given that all non-carcinogens tested gave negative results in all 5 transgenic models, providing evidence against “oversensitivity” concerns associated with such models.

<p>ILSI HESI ACT (2001) ILSI HESI Alternatives to carcinogenicity testing Workshop. Toxicologic Pathology 29 supplement: 1-351, Annex I.</p>	<p>The ILSI Alternative Methods for Carcinogenicity Testing (ACT) Workshop (held November 1–3, 2000 in Leesburg, Virginia, USA) featured two poster sessions to stimulate further discussion around the development, evaluation and application of alternative models for carcinogenicity assessment. In the Appendix I collected all the presented data.</p>
<p>Testing strategy Improvements in the Pharmaceuticals Sector</p>	
<p>'Putting Animal Welfare Principles and 3Rs into Action' - European Pharmaceutical Industry 2011 Report. http://www.efpia.eu/documents/14/61/Putting-animal-welfare-principles-and-3Rs-into-action-European-Pharmaceutical-Industry-2011-Report</p>	<p>The EFPIA (European Federation of Pharmaceutical Industries and Associations, reports on the state of the art on regulations, requirements, development, testing strategy, research and implementation of European Directives in the Pharmaceuticals Sector, in light of 3Rs principle.</p>
<p>Sistare FD, Morton D, Alden C, Christensen J, Keller D, Jonghe SD, et al., (2011). An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. Toxicologic Pathology 39 (4): 716-44.</p>	<p>Toxicology experts from the Pharmaceuticals Industry revised the carcinogenic potential of 182 marked and not marked products, demonstrating that the rat 2-years assay was not adding any valuable result or information to the cancer risk hazard of compounds which bear no clear histopathologic sign of rat tumour, any hormonal perturbation or positivity for genotoxicity. The authors advocate an update of regulatory requirements and guidelines for carcinogenicity studies.</p>

10. Reproductive toxicity – effects on fertility and developmental toxicity

Alexandra Rolaki, Malgorzata Nepelska, Susanne Bremer, Rabea Graepel, Anna Price & Andrew Worth

10.1 Introduction

Normal human reproduction and development depends on the timely execution of a wide range of processes including the formation, transport and release of gametes, fertilization, implantation, gestation, and, ultimately, the development of offspring that is eventually capable of successfully repeating the entire process. A comprehensive description of the reproductive functions and reproductive cycle can be found in text books on reproduction biology and/or reproductive toxicology (e.g. Gupta, 2011; Strauss & Barbieri, 2013).

In this document, the following definitions are taken from the Classification, Labelling and Packaging (CLP) regulation (EC, 200; Annex I: 3.7.1.1): a) “reproductive toxicity” includes adverse effects on sexual function and fertility in adult males and females, as well as developmental toxicity in the offspring; b) “effects on fertility” includes adverse effects on sexual function and fertility; and c) “developmental toxicity” includes adverse effects on development of the offspring.

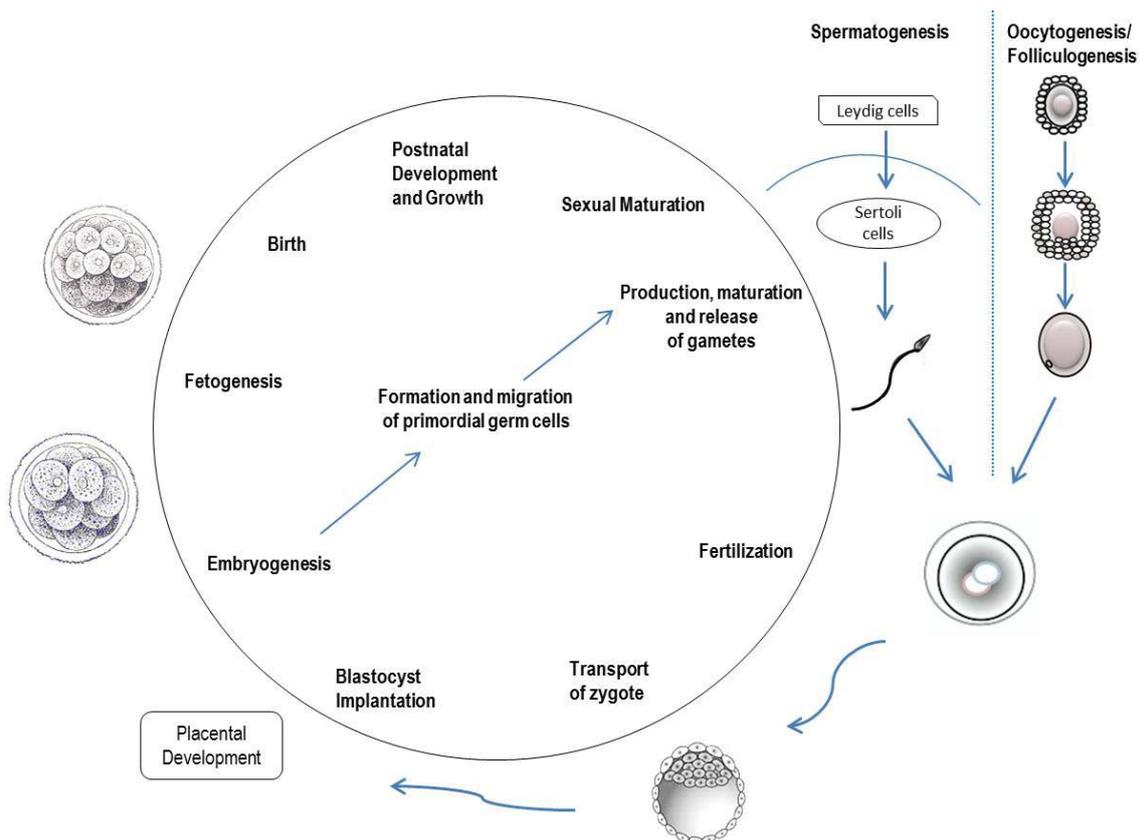


Figure 10.1. The main elements of the continuum of developmental stages and reproduction. Functional aspects such as libido, copulation, erection, ejaculation, oestrus cycle, pregnancy, labour and lactation are not included.

10.2 The traditional animal tests and their regulatory use

Various animal methods for testing effects on fertility and developmental toxicity are available and standardised as OECD Test Guidelines (see Table of References with Notes).

Under REACH, information on reproductive toxicity is required for chemicals with an annual production/importation volume of 10 metric tonnes or more. The standard information requirements are cumulative; requirements at higher tonnage levels add to the information requirements at lower tonnage levels. Standard information requirements include a screening study on reproduction toxicity (OECD TG 421/422) at Annex VIII (10-100 t.p.a), a prenatal developmental toxicity study (OECD 414) on a first species at Annex IX (100-1000 t.p.a), a two-generation reproductive toxicity study (OECD 416) and, if not conducted already at Annex IX, a prenatal developmental toxicity study on a second species at Annex X (≥ 1000 t.p.a.). Taken together, at Annex X level, information on reproduction toxicity could be available from a screening reproductive toxicity study, prenatal developmental toxicity studies on two species, two-generation reproductive toxicity study, as well as from repeated dose toxicity studies. In the REACH, there is an obligation to carry out vertebrate testing only as a last resort and to consider all other options before performing or requiring testing, as described by Articles 13(1) and 25(1) in the REACH legal text.

Under the Biocidal Products Regulation (BPR), which entered into force on 1 September 2013 (replacing the Biocidal Products Directive, 98/8/EEC), information is also required on reproductive toxicity for active substances as part of core data set and additional data set (EU 2012, ECHA 2013). As a core data set, prenatal developmental toxicity study (EU TM B.31) in rabbits as a first species and a two-generation reproduction toxicity study (EU TM B.31) are required. OECD TG 443 (Extended One-Generation Reproductive Toxicity Study) shall be considered as an alternative approach to the multi-generation study. The need of a further prenatal developmental toxicity study on second species (part of additional data set) may be decided based on the results of the first test and all available information.

The CLP regulation is based on the United Nations (UN) Globally Harmonized System (GHS) for classification and labelling. Also many other UN Member Countries are implementing (or will implement) GHS, having thus a comparable classification system as in the EU.

The classification of a substance as toxic for reproduction may have important implications, such as the need to obtain authorisation for marketing of the substance. According to the CLP regulation, substances may be classified either into category 1A, 1B, or 2. Category 1 comprises substances that are known to be toxic for reproduction, based on evidence in humans (Category 1A) or presumed to be human reproductive toxicants, based on evidence from animal experiments (Category 1B). Substances that are suspected to be human reproductive or developmental toxicants, but for which the evidence is not sufficiently convincing to place the substance in Category 1, are assigned to Category 2.

The designation of a substance as toxic for reproduction follows criteria that do not consider the dose level of the substance at which reproductive toxicity occur, as long as excessive generalized toxicity does not occur. The classification and labelling system is based on the intrinsic properties of the substance and gives indication about the hazards without consideration of exposure or the effective dose level and therefore, does not provide information on the actual risk of the chemical.

10.3 Mechanistic understanding of the endpoints

Normal human fertility and development are regulated throughout the entire life-cycle by a finely tuned system of coordinated signals that direct the activity of multiple interdependent target cells. This complex signalling process involves a vast number of steroid hormones, growth factors and other molecules, each of which is programmed to be delivered to its target tissue in an appropriate amount and in a pre-defined time window. A mechanistic understanding of effects on fertility and developmental toxicity therefore needs to account for how chemicals disrupt normal physiological pathways, and how these perturbations translate into effects at higher levels of biological organisation and across multiple time scales. Adverse outcomes at the organism level include dysmorphogenesis, growth retardation, lethality, functional and behavioural alterations, and impaired fertility. These adverse effects are not necessarily based on cell growth and death, but rather on more subtle changes, for example on differentiation and intercellular/intracellular signalling pathways.

Given the complexity of processes in reproduction, it is not surprising that our knowledge of the underlying mechanisms and their perturbations is incomplete. A great deal of fundamental knowledge is being generated in biology and genetics, but this is still a long way from being applied in toxicology.

10.4 Status of alternative methods

Given the diversity of physiological processes associated with the mammalian reproductive cycle, and the complexity of the underlying regulatory networks, it is not possible to model chemical effects on the whole cycle with a single or limited number of non-animal approaches.

The approach taken so far has been to break down the reproductive cycle into its main biological components, some of which are then studied individually with a view to improving mechanistic understanding and developing mechanistically-based *in vitro* tests. Indeed, a substantial amount of research has been conducted to develop non-animal approaches, and in particular *in vitro* tests. These efforts have been reviewed in detail elsewhere (Worth & Balls, 2002; Adler et al., 2011) and in a special issue of *Reproductive Toxicology* summarizing key results of the ReProTect project (Piersma et al., 2010). The relationship between the main *in vitro* tests developed, and the part of the reproductive cycle they are intended to mimic, is given in Table 10.1.

For some reproductive tissues and for some aspects of the reproductive cycle *in vitro* systems exist, but these have been used mainly for mechanistic research and have not been optimized for toxicity testing. Most of them consist of primary cultures or organ cultures.

Culture systems of pre-implantation embryos are well established because of interest in *in vitro* fertilization techniques and reproductive cloning, but little attention has been paid to these as toxicology models.

Stem cells provide a promising means of assessing chemical effects on stem cell differentiation into cells of the three germ layers as well as on their further differentiation and maturation into functional cells such as neurons or heart cells. In addition, stem cells possess the capacity to differentiate into any other cell type, which is valuable in the development of tissue-specific toxicity assays.

In the case of developmental neurotoxicity (DNT), a range of *in vitro* models exists that allow to study neurodevelopmental specific processes in a quantitative manner and could be included into a

testing strategy for initial screening and prioritisation of chemicals with concern for developmental neurotoxicity. As for development from the early embryonic period through adulthood, no models are available.

Despite the impressive number of *in vitro* methods that have been developed (Adler et al., 2011; Moschallski et al., 2011), most of them are not yet used in the regulatory assessment of reproductive toxicity. Many of the currently available assays have been developed for elucidating possible mechanisms of action, so they are not directly predictive of any specific adverse effects on reproduction. In addition, most of the assays do not consider the various aspects of absorption, distribution, metabolism and excretion (ADME), which need to be taken into account in determining the *in vivo* toxicological profile of a chemical. Another challenge consists in distinguishing general toxicity from reproductive toxicity, especially organ specific developmental effects, (even by means of whole embryo assays; see below). Furthermore, the influence of the maternal organism (maternal toxicity) on developmental toxicity cannot be directly captured by *in vitro* assays. The status of the main *in vitro* methods for reproductive toxicity testing is given in Table 10.2. In spite of the difficulties in developing non-animal approaches for reproductive toxicity, more approaches for developmental toxicity and the endocrine system are available and closer to regulatory applicability as supporting information than methods assessing effects on fertility.

Table 10.1. Relationship of available *in vitro* test methods to steps in the reproductive cycle

Step in reproductive cycle	<i>In Vitro</i> Test Method
Embryogenesis / Foetogenesis	Embryonic Stem Cell Test (EST)
Embryogenesis / Foetogenesis	The micromass test (MM)
Embryogenesis / Foetogenesis	The rodent whole embryo culture test (WEC)
Embryogenesis / Foetogenesis	The zebrafish embryo teratogenicity assay
Embryogenesis / Foetogenesis	Frog embryo teratogenesis assay xenopus (FETAX)
Embryogenesis / Foetogenesis	The chicken embryotoxicity screening test (CHEST)
Transport of substances through the placenta	Placental perfusion assay
Placental development / Blastocyst implantation	Trophoblast cell assay
Production and release of gametes	Leydig Cell Assay
Production and release of gametes	Sertoli 3-D cell system
Production and release of gametes / mutagenic effects	Comet and ReProComet Assay
Production and release of gametes	Computer Assisted Sperm Analysis (CASA)
Production and release of gametes	Follicle culture bioassay (FBA)
Oocyte Maturation	In vitro bovine oocyte maturation assay (bIVM)
Blastocyst implantation	Mouse Peri-Implantation Assay (MEPA)
Fertilization / Blastocyst implantation	In vitro bovine fertilization assay (bIVF)

10.4.1 *In vitro* methods for fertility

Some *in vitro* methods have been designed to investigate the cytotoxic effects of chemicals on the main components of the male and female reproductive organs and systems without functional fertility and reproductive performance. The status of the main *in vitro* methods is presented in Table 10.2.

In the male gonads, the three critical cell types sensitive to chemical exposure are: germ cells, Leydig and Sertoli cells. Given the importance of the quality and quantity of spermatozoa in male reproductive capability, the measurement of sperm parameters is one of the essential endpoints in assessing reproductive toxicity. Male reproductive toxicity is mainly assessed *in vitro* by Computer Assisted Sperm Analysis (CASA), the Leydig cell assay, the Sertoli cell assay and the ReProComet assay (Cordelli et al., 2007; La et al., 2010; Schleh & Leoni 2013).

In the female gonads, the essential functional and morphological unit is the follicle of the ovary. Each follicle consists of three basic cell types: theca and granulosa cells and the oocyte. A number of different *in vitro* assays have been developed to assess female reproductive toxicity, including the follicle culture bioassay (FBA), the bovine oocyte maturation assay (bIVM), the bovine fertilisation test (bIVF), and the mouse peri-implantation assay (MEPA) (Anon 2010; Lemeire et al., 2007; Lazzari et al., 2008; Luciano et al., 2010; Van Merris et al., 2007).

10.4.2 *In vitro* methods for developmental toxicity

In vitro methods for developmental toxicity testing are more numerous, and in general closer to regulatory applicability, than *in vitro* methods for assessing effects on fertility. The status of the main methods is given in Table 10.2.

The main *in vitro* methods for assessing placental transfer are the placental perfusion assay (Myren et al., 2007; Mose et al., 2008; Myllynen et al., 2009) and the trophoblast cell assay (Morck et al., 2010). The latter uses the immortalized BeWo trophoblastic cell line of human origin, while the former is an *ex vivo* placental perfusion. The placenta differs more than any other organ between species (Leiser & Kaufmann, 1994). This is the primary reason to develop models utilizing human tissue to study placental functions and to assess the placental transfer of compounds and their effects on placental formation and pregnancy. It should be noted however that the placental perfusion assay is relatively low throughput and difficult to perform.

Exposure to chemicals during development can affect nervous system development (as well as other systems) and cause transient or irreversible morphological and functional changes at doses much lower than those affecting adult brain function (Claudio et al., 2000; Tilson, 2000; Eriksson, 1997). A particular challenge in developmental neurotoxicity (DNT) testing is that outcome of exposure to potentially neurotoxic chemicals depends not only on dose and duration of exposure but also on the developmental stage at the time of exposure. Presently there are many *in vitro* test methods that allow evaluation of the effects on key neurodevelopmental processes including neuronal progenitor cell proliferation, migration, differentiation into neurons and glia, neuronal outgrowth, myelination, synaptogenesis and network formation (Hogberg et al., 2009; Fritsche et al., 2011; Bal-Price et al., 2012). However, human *in vitro* models that have been developed for DNT testing are highly recommended as the most relevant for prediction of human toxicity. Currently, it is possible to investigate toxicity induced by chemicals on key neurodevelopmental processes using human neuronal *in vitro* systems derived from progenitor- or human stem cells (embryonic, adult or induced pluripotent stem cells). They have been studied intensively over the past decade because they are primary, of human origin, and expandable although not immortalized (Buzanska et al., 2009). These human *in vitro* models allow to study most of the DNT processes including human neural progenitor cells commitment, proliferation, migration, differentiation, neurite outgrowth, synaptogenesis and network formation (Bal-Price et al., 2012; Fritsche et al., 2011).

The effects of chemicals on these DNT-specific processes can be studied *in vitro* in a quantitative manner to characterise the concentration-response relationship (Radio et al., 2008; Bal-Price et al., 2012; Balmer et al., 2012; Zimmer et al., 2012).

It is postulated that the developmental effects of pre- and/or post-natal exposure may be linked not only to behavioural changes of children (e.g. autism, attention deficit disorder, mental retardation) (Grandjean and Landrigan, 2006,) but also to adult cognitive and mental disorders such as Parkinsons (Fox et al., 2012). Therefore, due to the complexity of relationship between exposure and toxicity, DNT is also an area representing considerable challenge for predictive toxicology. DNT often manifests itself in behavioural and other disturbances that are hard to model *in vitro* (Van Thriel et al., 2012) but may be studied to some extent using non-mammalian models such as zebrafish (Padilla et al., 2011).

10.4.3 *Non-mammalian* models for developmental toxicity

In recent years, there has been growing interest in the use of non-mammalian organisms, and in particular zebrafish embryos, as a test system for developmental toxicity, since zebrafish embryo

development is very similar to embryogenesis in higher vertebrates including humans (Weigt et al., 2011; De Esch et al., 2012). Additionally, the zebrafish embryo possesses a set of convenient features including its small size, the rapid embryo development (major organs developed by 120 hours post-fertilization [hpf]), transparency of the embryo, large number of eggs available, full DNA sequence available, and an extensive supporting data- and literature- base (Yang et al., 2009). An important attribute of the zebrafish embryo, from a regulatory point of view, is the fact that its development up to the independently feeding larval forms does not require permission by responsible authorities (according to the EU directive 2010/63/EU on the protection of animals used for scientific purposes). However, Directive 2010/63/EU only sets the minimum requirements, and the individual EU Member States may set stricter requirements (see also Chapter 13). The exact onset of the independently feeding zebrafish is a matter of debate but it is clearly recommended that from 120 hours post fertilisation it requires permission to be used as a laboratory animal (Strahle et al., 2012). The zebrafish assay has been recommended as a screening tool for the prioritization of chemicals (Crofton et al., 2011).

10.4.4 ECVAM-validated in vitro tests for developmental toxicity

Three *in vitro* methods for embryotoxicity testing have been endorsed by the ECVAM Scientific Advisory Committee (ESAC) as scientifically validated (ECVAM, 2002): the mouse Embryonic Stem cell test (EST), the Micromass assay (MM), and the whole embryo culture (WEC). The EST and WEC were considered valid for distinguishing between non, weak/moderate and strong embryotoxicants, whereas the MM was considered valid for identifying strongly embryotoxic chemicals.

Of these three methods, the EST is the only method that does not entail the killing of animals for the tissue used in the assays. Although the test was formally validated additional investigations are needed to further define the relevance of the EST in the regulatory context. To date, the EST has not been accepted by the regulatory community, partly because of the unacceptably high false positive and negative prediction rates, and partly because of the need to improve the definition of the applicability domain, defined as the chemical classes and/or ranges of test method endpoints for which the model makes reliable predictions (Marx-Stoelting et al., 2009; Spielmann et al., 2006). An interesting development by researchers at Wageningen University is to combine the use of the EST with Physiologically Based Biokinetic (PBBK) modelling in order to derive quantitative points of departure *in vitro*, which are then extrapolated to *in vivo* points of departure for use in risk assessment (A. Punt, personal communication).

While these methods cannot on their own replace animal testing, they could be potentially used as supporting elements in weight of evidence or read-across approaches reducing animal use. It has been suggested that the value of these methods could be increased by incorporating molecular-based markers through the application of proteomic and toxicogenomic approaches (Piersma, 2006; van Dartel et al, 2010). The ESNATS project (see below) is exploring this possibility.

10.4.5 QSAR models for developmental toxicity

There have been a limited number of efforts aimed at developing Quantitative Structure-Activity Relationships (QSARs) for developmental toxicity, as reviewed elsewhere (Piparo & Worth, 2010; Cronin & Worth, 2008). Software tools based on QSARs include CAESAR, Derek, HazardExpert, MultiCASE, Leadscape, PASS and TOPKAT. Apart from the inherent difficulty of relating chemical structure to complex endpoints related to the reproductive cycle, modelling efforts have been limited by a lack of high quality datasets that can be used to train the models. In the case of developmental toxicity, two potentially useful databases in the public domain are the US EPA's

ToxRef Database (<http://www.epa.gov/NCCT/toxrefdb>) and the International Life Science Institute (ILSI) developmental toxicity database (<http://www.ilsi.org/Lists/Activities/AllItems.aspx>).

The only freely available software model for developmental toxicity is the CAESAR model (Cassano et al., 2010; <http://www.caesar-project.eu>), which is part of the VEGA suite. The same model is included in the US EPA's The Toxicity Estimation Software Tool (T.E.S.T.; <http://www.epa.gov/nrmrl/std/qsar/qsar.html>). This model was built on a data set of 292 compounds (Arena et al., 2004) by using a random forest modelling approach and with reference to the US FDA classification system which has five categories: category A means negative human studies; category B means negative animal studies and no human studies, or positive animal studies and negative human studies; category C means positive animal studies and no human studies, or no studies at all; category D means positive human studies; and category X means animal or human studies show abnormalities and/or evidence of foetal risk. To build the model, a compound was considered non-developmentally toxic if it had been classified under the FDA scheme in categories A or B, and was considered developmentally toxic if it had been classified to the categories C, D, or X. The usefulness of this model for the purposes of REACH and CLP has not been evaluated.

In a study carried out by the JRC for EFSA (Worth et al., 2011), the predictive abilities of six models (Derek, CAESAR, TOPKAT, Leadscope, HazardExpert, and PASS) were evaluated against a dataset of 76 pesticides, an extended dataset of 135 chemicals including the 76 pesticides and 69 additional chemicals, and a dataset of 366 pesticides from the US EPA's ToxRef database. None of the models, and no two-model battery of models, was found useful in predicting the absence of developmental toxicity, since the negative predictivities were less than 50%. However, some QSAR tools, such as Derek, HazardExpert and PASS, might be useful for the identification of potential developmental toxicants, due to their high positive predictivities of 81-96%. A possible approach, demonstrated in the JRC report, is to use a stepwise assessment strategy in which QSARs are used in a first step to identify positives, followed by the use of read-across in a second step to identify negatives. The feasibility of this approach would depend on the availability of reliable data for analogues of the chemical of interest.

With a view to developing structure-based prediction models, there have been recent efforts to better understand the relationship between chemical (sub)structures and toxicity pathways associated with reproductive toxicity. For example, based on a detailed review of 716 chemicals, Wu et al. (2013) have identified 25 structural groups including 5 groups associated with specific receptor-mediated interactions, and 20 more general chemical groups. The underlying structural alerts were arranged into a decision tree. Each step in the decision tree leads to another step or a final prediction (yes/no for reproductive toxicity effects). The decision tree is proposed for use in priority setting and in informing weight-of-evidence assessments.

Another recent approach, which also provides a means of developing QSARs and grouping chemicals for read-across, is to use chemotypes as chemical descriptors. These are essentially structural alerts (SAs), but in contrast to traditional SAs, they encode information on atom-specific and whole-molecule physicochemical properties, providing a great capacity to identify reactivity and toxicity-informed chemical features. Preliminary studies have demonstrated the usefulness of chemotypes in predicting developmental toxicity potential (cleft lip and palate formation), as illustrated in a poster presentation at the 2013 SOT conference (Yang et al., 2013). Chemotypes have been identified for molecular initiating events in various signalling pathways whose disruption can lead to such developmental effects, including the hedgehog, glucocorticoid, RXR- α , Wnt, SHH, folate, AhR and sterol biosynthesis pathways (C. Yang, personal communication). A software tool, the Chemotyper, is freely available from Altamira (<http://www.altamira-llc.com>), for the screening

of datasets against a predefined set of 686 chemotypes (including chemotypes for developmental toxicity and other adverse outcomes).

10.4.6 Pathway-based approaches and systems biology (mathematical models)

Various mammalian and non-mammalian organisms are being explored for their usefulness as screening systems based on the perturbation of gene regulatory pathways by reproductive, and in particular developmental, toxicants.

Pathway-based endpoints observed in mammalian cells can provide a means of establishing points of departure and enabling *in vitro* to *in vivo* extrapolation (Anderson et al., 2005). For example, using *in vitro* data on aromatase activity and mRNA levels in rat granulosa cells, Quignot and Bois developed a mathematical (systems biology) model that simulates the synthesis and secretion of steroid hormones in the rat ovary and the effects on these processes of endocrine active chemicals (Quignot & Bois, 2013).

The use of pathway-based endpoints in non-mammalian cells is based on the fact that well-defined and conserved pathways involved in development are conserved across species and found in model organisms such as *C. elegans* (Goussen et al, 2013), *Drosophila* (Uysal et al., 2013), the zebrafish embryo (Perkins et al, 2013), and the sea urchin (Davidson et al., 2002; Ben-Tabou de-Leon & Davidson, 2007). For example, the retinol metabolism pathway is highly conserved across vertebrates, and perturbations to this pathway have been causally linked to developmental effects (skeletal deformities) in both zebrafish embryos and mammals (Laue et al, 2011). On this basis, Perkins et al (2013) proposed a pathway-based point-of-departure for these effects based on transcriptomic changes in the zebrafish embryo. They showed, for example, that the pathway based lowest observed effect level (LOEL) (2.8 μ M) for flusilazole-induced developmental effects in zebrafish embryo was ten-fold lower than the LOEL based on apical effects (>28 μ M). Key events (upstream of the apical effect) have been widely used as conservative points of departure (Hill et al, 1998).

10.4.7 Adverse Outcome Pathways for developmental toxicity

The adverse outcome pathway (AOP) approach is also being explored as a basis for describing and predicting developmental toxicity. Examples include an AOP for paraoxon-induced effects on zebrafish embryogenesis (paraoxon is a reference compound for AChE inhibition) (Yozzo et al., 2013), and an AOP linking the disruption of embryonic vascular development (for example by thalidomide, oestrogens, endothelins, dioxin, retinoids, cigarette smoke, and metals) to developmental toxicity (Knudsen & Kleinstreuer, 2011).

10.4.8 Virtual embryo to model developmental toxicity

Morphogenesis is a complex tissue phenomenon consisting of multiple processes that are controlled by the genome and affected by the environment (including chemicals and maternal-foetal interactions). The successful modelling of chemically-induced dysmorphogenesis needs to describe multiple cell states (proliferation, differentiation, death, motility, shape, adhesivity) as well as spatiotemporal heterogeneities in cell viability and metabolic demands. Furthermore, a successful model needs to predict the cellular responses to genetic (programmed) and environmental signals, and the emergent morphogenetic properties associated with collective cellular behaviour. This represents a considerable scientific and technical challenge.

The US EPA's Virtual Embryo (v-Embryo™) research project (<http://epa.gov/ncct/v-Embryo/index.html>) is applying the techniques of computational systems biology to explore the

feasibility of developing a mathematical model that is able to accurately predict the potential for environmental chemicals to affect the embryo. The initial focus of v-Embryo™ is to simulate early eye, vascular and limb development (Knudsen et al, 2011). The modelling is supported by data-generating experiments using stem cells and zebrafish.

10.4.9 EU research projects for reproductive toxicity

Through its framework programme (FP) for research and development, the European Commission has invested substantially in developing alternative methods to animal testing for reproductive toxicity. In particular, three main projects have been funded.

The FP6 FunGenES (Functional Genomics in Engineered ES cells) (2004-2007) project contributed to characterisation of the molecular mechanisms underlying differentiation of embryonic stem cells to somatic cells, which constitutes the basis for further use of these cells for toxicity testing. While this was a toxicology project not directly focused on developmental toxicity, it developed an understanding of embryonic stem cell self-renewal, differentiation and lineage commitment to different organ-specific cells, including the identification of target genes for therapeutic intervention. In addition, new cellular and molecular tools were developed to characterise gene function in tissue-specific cell populations (functional genomics), along with new embryonic stem cell-based approaches to screening of small candidate molecules for therapeutic applications in human diseases.

The FP6 ReProTect project (2004-2009) consisted of a consortium of 32 EU organizations with the aim of developing and optimising *in vitro* test methods in the area of reproductive toxicology (<http://www.reprotect.eu>). The project made advances on 15 *in vitro* tests that target adverse effects on mammalian female and male fertility and reflect various toxicological mechanisms such as effects on Leydig and Sertoli cells, folliculogenesis, germ cell maturation, sperm cell motility, steroidogenesis, the endocrine system, fertilisation, and on the pre-implantation embryo. The project aimed to define better the applicability domain of the validated embryonic stem cell test by testing additional substances.. Proteome analyses performed on mouse and human embryonic stem cells treated with selected chemicals during neuronal and cardiac differentiation were conducted in order to select biomarkers with a higher predictivity. In parallel, an integration of a metabolic system into the embryonic stem cell test (EST) was put forward. A special issue of Reproductive Toxicology summarises key results of the ReProTect project (Piersma et al., 2010).

The development/optimisation of each test was performed according to the ECVAM modular approach and the INVITTOX protocols of the assays were made available. Besides, an independent statistical evaluation of the raw data has evaluated the reproducibility of the developed tests. The EST is now being challenged by testing newly selected compounds and using the previously developed standard operating procedures and new read-out technologies.

An ECVAM/ReProTect workshop was organised in 2008 to discuss the status and future of the various derivatives of the EST assay, in order to identify how much flexibility can be introduced to the *in vitro* tests based on embryonic stem cells and their applicability for regulatory decision making. The outcome of the workshop was published in 2009 (Marx-Stoelting et al., 2009), in which conclusions and recommendations were made concerning the need to add a metabolic system to the EST, and the possibility to add new endpoints, including additional cell differentiation endpoints (in non-cardiac tissues), and new molecular markers (using omics and reporter gene assays) for toxic effects on embryonic development.

The FP7 ESNATS (Embryonic Stem cell-based Novel Alternative Testing Strategies) project (2007-2013) is developing a battery of toxicity tests based on embryonic stem cells (ESCs), in particular human ESC (hESCs), with a view to accelerating drug development and reducing related R&D costs (<http://www.esnats.eu>). ESNATS has established *in vitro* systems that recapitulate different critical periods of human early neuronal development.

The FP7 ChemScreen (Chemical substance *in vitro/in silico* screening system to predict human- and ecotoxicological effects) project (2010-2013) aims to generate a simple, rapid screening system for reproductive toxic chemicals, capable of being performed within the timeframe of the REACH implementation (<http://www.chemscreen.eu>). Unlike other EU-funded projects, ChemScreen places more emphasis to the development and integrated use of *in silico* and *in vitro* methods as a means of identifying potential reproductive toxicants. A High Throughput Screening (HTS) panel comprises *in vitro* assays expressing reporter genes associated with a wide range of signalling pathways potentially involved in reproductive toxicity. In addition to these HTS methods, assays are being established in more complex, differentiated cells such as murine ES cells by introducing reporter systems for signalling pathways controlling key differentiation pathways in embryonic development (Wnt/, SHH-, TGF β -, Delta/Notch-, and the RTK-signalling pathway). Furthermore, PBBK models are being developed to translate *in vitro* concentrations to *in vivo* dose levels.

10.5 Conclusions

Information on reproductive toxicity is required for the purposes of REACH, the Biocidal Products Regulation, Plant Protection Products Regulation and the CLP regulation. Since the assessment of these endpoints requires a considerable number of animals and testing costs are high, it is a priority area for the development of alternative methods. However, it is also one of the most challenging from both scientific and technical points of view.

Currently available *in vitro* models cover only a small part of the reproductive cycle, with many steps not being covered at all. These *in vitro* tests range from cell-based models, including primary cultures and stem cells, to organ cultures, and whole embryo cultures. Recent work has also highlighted the value of non-mammalian species as model organisms, and in particular zebrafish embryos.

A few QSAR models have been developed for endpoints related to developmental toxicity, but these do not provide information that is directly applicable in regulatory decision making. The recent use of chemotypes in developing mechanistic QSARs should prove useful in the screening of chemical inventories for specific positive effects (such as the potential for cleft palate formation).

In conclusion, currently available *in vitro* tests and QSAR models may be useful for screening chemical inventories and identifying positives (priority setting). In addition, they may be useful elements for supporting chemical grouping and read-across, and weight-of-evidence assessments. However, these latter applications need to be evaluated on a case-by-case basis.

The challenge in replacing animal tests for the effects on fertility and developmental toxicity is that many possible organs and physiological processes can be affected, and knowledge of the underlying regulatory networks and the possible impacts of their perturbation is limited. However, increasing the body of knowledge about key events of the reproductive cycle leading to adverse effects is expected to provide a basis for understanding the multiple disruption mechanisms and their interactions, and for translating this knowledge into test systems at least for some parts of the

reproductive cycle. In reproductive toxicity it is challenging to capture in a test system even the most relevant separate and interrelated endpoints.

The current approach of developing *in vitro* tests to mimic some key elements of the reproductive cycle has led to the development of many potentially useful “building blocks”, but these have still a number of limitations. For example, the lack of metabolic competence in many *in vitro* systems needs to be compensated by addition of metabolizing systems. Other limitations reflect the absence of whole animal physiology and cannot be captured by individual test methods, but rather by an integrated approach. For example, for the purposes of risk assessment, a range of biokinetic (ADME) parameters (in addition to metabolism) need to be measured or computed, and integrated by means of PBBK models. This may allow toxicologically relevant points of departure to be derived from *in vitro* concentration-response curves to be extrapolated from *in vitro* to *in vivo*.

Integrated approaches testing and assessment will need to integrate both *in vitro* (in particular cell-based assays) and computational tools (in particular QSARs and PBBK models). Based on the current state of the art, it is not anticipated that such testing strategies will completely replace the need to assess reproductive toxicity *in vivo*. However, they could provide a means to reduce animal testing.

Finally, the premise that it is sufficient to mimic key parts of the reproductive cycle can be questioned. In other words, the toxic effects of chemicals on reproduction are more than the sum of multiple cellular effects.

In the longer-term, the increased use of mechanistically-based *in vitro* tests, for example cell-based assays expressing signalling pathways associated with reproductive processes, along with the development of systems biology-based mathematical models to perform *in vitro* to *in vivo* extrapolations, may provide reliable means supporting risk assessment and classification and labelling.

10.6 References

- Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhrer S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A & Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85: 367-485
- Arena VC, Sussman NB, Mazumdar S, Yu S & Macina OT (2004). The utility of Structure-activity relationship (SAR) models for prediction and covariate selection in developmental toxicity: comparative analysis of logistic regression and decision tree models. *SAR and QSAR in Environmental Research* 15: 1-18.
- Andersen ME, Thomas RS, Gaido KW & Conolly RB (2005). Dose-response modeling in reproductive toxicology in the systems biology era. *Reproductive Toxicology* 19: 327-337.
- Balmer N V, Weng MK, Zimmer B, Ivanova VN, Chambers SM, Nikolaeva E, Jagtap S, Sachinidis A, Hescheler J, Waldmann T & Leist M (2012). Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome. *Human Molecular Genetics* 21: 4104-4114.
- Bal-Price AK, Coecke S, Costa L, Crofton KM, Fritsche E, Goldberg A, Grandjean P, Lein PJ, Li A, Lucchini R, Mundy WR, Padilla S, Persico AM, Seiler AEM & Kreysa J (2012). Advancing the science of developmental neurotoxicity (DNT): testing for better safety evaluation. *ALTEX* 29: 202-215.
- Ben-Tabou de-Leon S & Davidson EH (2007). Gene regulation: gene control network in development. *Annual Review of Biophysics and Biomolecular Structure* 36: 191-212.
- Bremer S, Balduzzi D, Cortvrindt R, Daston G, Eletti B, Galli A, Huhtaniemi I, Laws S, Lazzari G, Liminga U & Coecke S (2005). The Effects of Chemicals on Mammalian Fertility. *ATLA* 33: 391-416.
- Buzanska L, Sypecka J, Nerini-Molteni S, Compagnoni A, Hogberg HT, del Torchio R, Domanska-Janik K, Zimmer J & Coecke S (2009). A human stem cell-based model for identifying adverse effects of organic and inorganic chemicals on the developing nervous system. *Stem Cells* 27: 2591-601.
- Claudio L, Kwa WC, Russell AL & Wallinga D. (2000). Testing methods for developmental neurotoxicity of environmental chemicals. *Toxicology and Applied Pharmacology* 164: 1-14.
- Crofton KM, William MR, Lein PJ, Bal-Price A, Coecke S, Seiler AEM, Kanut H, Buzanska L & Goldberg A (2011). Developmental Neurotoxicity Testing: Recommendations for Developing Alternative Methods for the Screening and Prioritization of Chemicals. *ALTEX* 28: 1-11.
- Cronin MTD & Worth AP (2008). (Q)SARs for Predicting Effects Relating to Reproductive Toxicity. *QSAR & Combinatorial Science* 27: 91-100.
- Davidson EH et al. (2002). A genomic regulatory network for development. *Science* 295: 1669-1678.
- De Esch C, Slieker R, Wolterbeek A, Woutersen R & de Groot D (2012). Zebrafish as potential model for developmental neurotoxicity testing: a mini review. *Neurotoxicology and Teratology* 34: 545-553.

EC (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353.

ECHA (2013). Guidance on Regulation (EU) No 528/2012 concerning the making available on the market and use of biocidal products (BPR) Version 1.0.

Eriksson, P. (1997). Developmental neurotoxicity of environmental agents in the neonate. *Neurotoxicology* 18, 719-726.

ECVAM (2002) ECVAM Scientific Advisory Committee (ESAC) statements on alternative methods. [http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions/?searchterm=ESAC Statements](http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/scientific-advice-stakeholders-networks/ecvam-scientific-advisory-committee-esac/statements-opinions/?searchterm=ESAC%20Statements)

EU (2012). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Commission Delegated Regulation (EU) No 736/2013 and (EU) No 837/2013 amending Regulation (EU) No 528/2012.

Flint OP & Orton TC (1984). An in vitro assay for teratogens with cultures of rat embryo midbrain and limb bud cells. *Toxicology and Applied Pharmacology* 76: 383–395.

Fox DA, Grandjean P, de Groot D, Paule MG (2012). Developmental origins of adult diseases and neurotoxicity: epidemiological and experimental studies. *Neurotoxicology* 33:810-816.

Fritsche, E., Gassmann, K., and Schreiber, T. (2011). Neurospheres as a model for developmental neurotoxicity testing. *Methods Mol. Biol.* 758, 99-114.

Goussen B, Parisot F, Beaudouin R, Dutilleul M, Buisset-Goussen A, Péry ARR & Bonzom J-M (2013). Consequences of a multi-generation exposure to uranium on *Caenorhabditis elegans* life parameters and sensitivity. *Ecotoxicology* 22: 869-878.

Grandjean, P. and Landrigan, P. J. (2006). Developmental neurotoxicity of industrial chemicals. *Lancet* 368, 2167-2178.

Gupta RC, Ed (2011). *Reproductive and Developmental Toxicology*. Academic Press, Amsterdam.

Hill RN, Crisp TM, Hurley PM, Rosenthal SL & Singh DV (1998). Risk assessment of thyroid follicular cell tumors. *Environmental Health Perspectives* 106: 447–457.

Hogberg HT, Kinsner-Ovaskainen A, Hartung T, Coecke S & Bal-Price AK (2009). Gene expression as a sensitive endpoint to evaluate cell differentiation and maturation of the developing central nervous system in primary cultures of rat cerebellar granule cells (CGCs) exposed to pesticides. *Toxicology and Applied Pharmacology* 235: 268–286.

Krug AK, Balmer N V, Matt F, Schönenberger F, Merhof D & Leist M (2013). Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. *Archives of Toxicology* 87: 2215-2231.

Knudsen TB & Kleinstreuer NC (2011). Disruption of embryonic vascular development in predictive toxicology. *Birth Defects Research Part C, Embryo Today : Reviews* 93: 312–323.

Laue K, Pogoda H-M, Daniel PB, van Haeringen A, Alanay Y, Ameln von S, et al. (2011). Craniosynostosis and Multiple Skeletal Anomalies in Humans and Zebrafish Result from a Defect in the Localized Degradation of Retinoic Acid. *Am J Hum Genet* 89:595–606.

Li H, Van Ravenzwaay B, Rietjens IMCM & Louisse J (2013). Assessment of an in vitro transport model using BeWo b30 cells to predict placental transfer of compounds. *Archives of Toxicology* 87: 1661-1669..

Lo Piparo E & Worth A (2010). Review of QSAR Models and Software Tools for predicting Developmental and Reproductive Toxicity. JRC report EUR 24522 EN. Publications Office of the European Union. Available from:

<http://publications.jrc.ec.europa.eu/repository>

Marx-Stoelting P et al. (2009). A review of the implementation of the embryonic stem cell test (EST). The report and recommendations of an ECVAM/ReProTect Workshop. *Alternatives to Laboratory Animals* 37: 313–328.

Moschallski N, Schuhmacher-Wolz U, Hassauer M & Schneider K (2011). Thematic Review: Alternative Methods for Reproductive Toxicity testing - Revision and update of the sector Reproductive Toxicity testing of the publicly available “DataBase service on ALternative Methods to animal experimentation (DB-ALM)”. Final Progress Report FoBiG , Forschungs- und Beratungsinstitut Gefahrstoffe GmbH reviewed and approved by EC JRC. Available from:

http://ec.europa.eu/enterprise/epaa/2_activities/2_3_comm_and_dissem/report_alternative_methods_reproductive_toxicity_testing.pdf

Myren M, Mose T, Mathiesen L & Knudsen LE (2007). The human placenta--an alternative for studying foetal exposure. *Toxicology in Vitro*: 1332-1340.

Padilla, S. and MacPhail, R. C. (2011). Using zebrafish to assess developmental neurotoxicity. In R.C. Gupta (ed.). *Reproductive and Developmental Toxicology*. London: Academic Press, 179-192.

Perkins EJ, Ankley GT, Crofton KM, Garcia-Reyero N, Lalone CA, Johnson MS, Tietge JE, Villeneuve DL (2013). Current perspectives on the use of alternative species in human health and ecological hazard assessments. *Environmental Health Perspectives* 121:1002-1010.

Piersma AH (2006). Alternative methods for developmental toxicity testing. *Basic & Clinical Pharmacology & Toxicology* 98: 427–431.

Piersma AH, Genschow E, Verhoef A, Spanjersberg MQI, Brown NA, Brady M, Burns A, Clemann N, Seiler A & Spielmann H (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Alternatives to Laboratory Animals* 32: 275–307.

Piersma AH et al (2010). ReProTect Special Issue. *Reproductive Toxicology* 30 (1): 1-218.

Quignot N & Bois FY (2013). A computational model to predict rat ovarian steroid secretion from in vitro experiments with endocrine disruptors. ed. Porollo A. *PloS One* 8: e53891.

Radio NM, Breier JM, Shafer TJ & Mundy WR (2008). Assessment of chemical effects on neurite outgrowth in PC12 cells using high content screening. *Toxicological Sciences* 105: 106–118.

Robinson JF & Piersma AH (2013). Toxicogenomic approaches in developmental toxicology testing. *Methods in Molecular Biology* 947: 451–473.

Robinson JF, Verhoef A, Pennings JL a, Pronk TE & Piersma AH (2012). A comparison of gene expression responses in rat whole embryo culture and in vivo: time-dependent retinoic acid-induced teratogenic response. *Toxicological Sciences* 126: 242–254.

- Rubinstein AL (2006). Zebrafish assays for drug toxicity screening. *Expert Opinion on Drug Metabolism & Toxicology* 2: 231–240.
- Schenk B, Weimer M, Bremer S, Van der Burg B, Cortvrindt R, Freyberger A, Lazzari G, Pellizzer C, Piersma A, Schäfer WR, Seiler A, Witters H & Schwarz M (2010). The ReProTect Feasibility Study, a novel comprehensive in vitro approach to detect reproductive toxicants. *Reproductive Toxicology* 30: 200–218.
- Spielmann H, Seiler A, Bremer S, Hareng L, Hartung T, Ahr H, Faustman E, Haas U, Moffat GJ, Nau H, Vanparys P, Piersma A, Sintes JR & Stuart J (2006). The practical application of three validated in vitro embryotoxicity tests. The report and recommendations of an ECVAM/ZEBET workshop (ECVAM workshop 57). *Alternatives to Laboratory Animals* 34: 527–538.
- Strähle U, Scholz S, Geisler R, Greiner P, Hollert H, Rastegar S, Schumacher A, Selderslaghs I, Weiss C, Witters H & Braunbeck T (2012). Zebrafish embryos as an alternative to animal experiments – A commentary on the definition of the onset of protected life stages in animal welfare regulation. *Reproductive Toxicology* 33: 128-132
- Strauss III JF & Barbieri RL, Eds (2013). *Yen & Jaffe's Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management*. Elsevier Health Sciences.
- Theunissen PT, Robinson JF, Pennings JL a, De Jong E, Claessen SMH, Kleinjans JCS & Piersma AH (2012). Transcriptomic concentration-response evaluation of valproic acid, cyproconazole, and hexaconazole in the neural embryonic stem cell test (ESTn). *Toxicological Sciences* 125: 430–438.
- Theunissen PT, Schulpen SHW, Van Dartel DAM, Hermsen SAB, Van Schooten FJ & Piersma AH (2010). An abbreviated protocol for multilineage neural differentiation of murine embryonic stem cells and its perturbation by methyl mercury. *Reproductive Toxicology* 29: 383–392.
- Tilson HA. (2000). Neurotoxicology risk assessment guidelines: developmental neurotoxicology. *Neurotoxicology* 21, 189-194.
- UN (2013) Globally Harmonized System of Classification and Labelling of Chemicals (GHS), 5th revised edition. United Nations, New York and Geneva, 2013
http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf.
- Uysal H, Semerdoken S, Colak DA & Ayar A (2013). The hazardous effects of the three natural food dyes on developmental stages and longevity of *Drosophila melanogaster*. *Toxicology and Industrial Health*. Epub ahead of print, 1 March 2013. doi: 10.1177/0748233713480206.
- Van Dartel DAM, Pennings JL, Van Schooten FJ & Piersma AH (2010). Transcriptomics-based identification of developmental toxicants through their interference with cardiomyocyte differentiation of embryonic stem cells. *Toxicology and Applied Pharmacology* 243: 420–428.
- Van Dartel DAM & Piersma AH (2011). The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. *Reproductive Toxicology* 32: 235–244.
- Van der Burg B, Kroese ED & Piersma AH (2011). Towards a pragmatic alternative testing strategy for the detection of reproductive toxicants. *Reproductive Toxicology* 31: 558–561.
- Van Thriel C, Westerink RHS, Beste C, Bale AS, Lein PJ & Leist M (2012). Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. *Neurotoxicology* 33: 911–924.

Weigt S, Huebler N, Strecker R, Braunbeck T & Broschard TH (2011). Zebrafish (*Danio rerio*) embryos as a model for testing proteratogens. *Toxicology* 281: 25–36.

Worth AP & Balls M, eds. (2002). *Alternative (non-animal) methods for chemicals testing: current status and future prospects*. A report prepared by ECVAM and the ECVAM Working Group on Chemicals. *Alternatives to Laboratory Animals* 30: Supplement 1, July 2002.

Worth A, Fuart-Gatnik M, Lapenna S & Serafimova R (2011). Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Report produced for EFSA. Available from: <http://www.efsa.europa.eu/en/supporting/pub/169e.htm>

Wu S, Fisher J, Naciff JM, Laufersweiler MC, Lester C, Daston G & Blackburn K (2013). A framework for identifying chemicals with structural features associated with potential to act as developmental of reproductive toxicants. *Chemical Research in Toxicology* 26: 1840-1861.

Yang C, Arvidson K, Richard A, Worth A, Tarkhov A, Ringeissen S, Marusczyk J, Gasteiger J., Rathman J, Schwab C (2013). Chemotypes and Chemotyper: a new structure representation standard to include atomic/bond properties into structural alerts for toxicity effects and mechanisms. Poster presentation. Society of Toxicology, San Antonio, Texas, USA, 2013.

Yang L, Yu Ho N, Rudiger A, Legradi J, Weiss C, Reischl M, Mikut R, Liebel U, Muller F & Strahle U (2009). Zebrafish embryos as models for embryotoxic and teratological effects of chemicals. *Reproductive Toxicology* 28: 245-253

Yozzo KL, McGee SP & Volz DC (2013). Adverse outcome pathways during zebrafish embryogenesis: a case study with paraoxon. *Aquatic Toxicology* 126: 346–354.

Zhang C, Cao J, Kenyon JR, Panzica-kelly JM, Gong L & Augustine-Rauch K (2012). Development of a Streamlined Rat Whole Embryo Culture Assay for Classifying Teratogenic Potential of Pharmaceutical Compounds. *Toxicological Sciences* 127: 535–546.

Zimmer B, Lee G, Balmer N V, Meganathan K, Sachinidis A, Studer L & Leist M (2012). Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. *Environmental Health Perspectives* 120: 1116–1122.

Table 10.2. Alternative methods for reproductive toxicity testing

Method	Test system	Description	Applicability	Status
Embryonic stem cell test (EST)	Embryonic stem cell line	<p>Based on the assumption that chemicals with embryotoxic potential block the spontaneous development of embryonic stem cells into beating cardiac muscle cells within the first 10 days of embryonic development.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • cell differentiation • cell viability • cell proliferation 	<p>Suitable for screening the potential embryotoxic effects of substances, classifying them into three major classes:</p> <ul style="list-style-type: none"> • non-embryotoxic; • weakly embryotoxic; and • strongly embryotoxic. 	ESAC statement (ECVAM, 2002)
Micromass test (MM)	Limb bud cells and/or neuronal cells	<p>The cells are isolated from the limb or the cephalic tissues of mid-organogenesis embryos. The viability and differentiation after exposure to test chemicals is analysed.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • cell differentiation • cell viability • cell number • cell growth 	<p>Suitable for screening the potential embryotoxic effects of substances, classifying them into three major classes:</p> <ul style="list-style-type: none"> • non-embryotoxic; • weakly embryotoxic; and • strongly embryotoxic <p>Additionally, the method is suggested for the monitoring of development within a period and/or within defined organ system.</p>	ESAC statement (ECVAM, 2002)

Method	Test system	Description	Applicability	Status
Rodent post-implantation whole embryo culture (WEC)	Rat embryos	<p>Rat embryos are cultured during organogenesis <i>in vitro</i> and treated with test chemicals.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Embryo Morphology • Embryo Functionalities • Embryo Growth • Cell viability 	Suitable for identifying strongly embryotoxic or compounds causing malformations during the organogenesis in a complete mammalian embryo (an intact embryo is exposed <i>in vitro</i> , allowing the study of malformations).	ESAC statement (ECVAM, 2002)
Zebrafish embryo teratogenicity assay	Fertilised zebrafish (<i>Danio rerio</i>) eggs	<p>The method uses the fertilised fish eggs that are exposed to different concentrations of a test substance. At different time points, the exposed developing fish embryos are observed and scored for lethal, embryotoxic and/or teratogenic effects.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • lethality • malformations • growth retardation 	The method is used as a screening tool for potency to induce malformations, but also as a means of investigating specific mechanisms related to the induced malformations with certain substances.	Optimisation
Frog embryo teratogenesis assay xenopus (FETAX)	Whole frog (<i>Xenopus laevis</i>) embryo	<p>FETAX is a whole embryo screening assay, based on frog <i>Xenopus laevis</i>, to identify substances (assessed at different time points) that may pose a developmental hazard in humans.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • lethality • growth • retardation • malformations 	FETAX encompasses organogenesis and does not include later events of development.	Optimisation

Method	Test system	Description	Applicability	Status
Chicken embryotoxicity screening test (CHEST)	Chicken embryo	The chicken embryotoxicity screening test determines the toxic dose range in very early administration time (24 h) and the teratogenic dose range and covers late effects on the embryo development (days 2, 3 and 4). Endpoints: <ul style="list-style-type: none"> • mortality, • malformations, • embryo development, • blood vessel development • blood vessel coloration. 	Routine embryotoxicity screening and mechanistic studies	Optimisation
Placental perfusion assay	Human placenta	Human <i>ex vivo</i> placental perfusion which is set up 30 min. following the birth. To assess the transfer of substances from the maternal to the foetal side of the syncytiotrophoblast and to investigate molecular mechanisms of genotoxic compounds. Endpoints: <ul style="list-style-type: none"> • Transplacental transfer rate of compounds 	It serves to investigate the risks of foetal exposure to chemical compounds during pregnancy	Optimisation
Trophoblast cell assay	BeWo human trophoblast cell line	An assay that uses BeWo human trophoblast cell line, which have preserved many of the trophoblastic properties. It is considered as an appropriate <i>in vitro</i> system to investigate placenta transport. Endpoints: <ul style="list-style-type: none"> • Transplacental transfer rate of compounds 	It evaluates the placental transport of compounds during implantation and pregnancy.	R&D
Leydig Cell Assay	Leydig cell line BLT1-L17	The Leydig cell line BLT1-L17 has been developed to assess the disturbance of the endocrine system due to effects of chemicals. Endpoints: <ul style="list-style-type: none"> • Viability 	Toxicity assessment during spermatozoa development, with possible implications on fertility.	R&D

Method	Test system	Description	Applicability	Status
		<ul style="list-style-type: none"> • Growth • Testosterone production 		
Sertoli 3-D cell system	Combination of sertoli and germ cells	<p><i>In vitro</i> 3-D test model that replicates the composition, organization and functions of <i>in vivo</i> rat blood-testis barrier. The assay uses a combination of sertoli and germ cells.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Viability • Growth • Inhibin B production 	Hazard identification that might lead to fertility impairments	R&D
Comet and ReProComet Assay	Sperm cells	<p>The sperm cells are incubated with the test compound for 2 hours followed by flow cytometric analysis.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Viability • Morphology • DNA fractions 	For the detection of chemically induced DNA damage in bull sperm with possible implications on male fertility	Optimisation
Computer Assisted Sperm Analysis (CASA)	Spermatozoa	<p>Automated method to visualize and and evaluate consecutive images of viable sperms to obtain precise and valid information on the kinematics of individual sperms.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Spermatozoa viability, motility, velocity, motion, morphology • Metabolic effects 	Monitor effects of chemicals on spermatozoa with possible implications on fertility	Optimisation
Follicle culture bioassay (FBA)	Mouse pre-antral follicles	Mechanically isolated mouse pre-antral follicles are cultured <i>in vitro</i> until the pre-ovulatory phase followed by <i>in vitro</i> ovulation and maturation of the	Evaluation of chemicals during folliculogenesis, steroidogenesis and	Optimisation

Method	Test system	Description	Applicability	Status
		<p>oocyte. During the growth period the follicle is exposed to chemicals in an acute or chronic manner.</p> <p>Endpoints:</p> <p>Follicle:</p> <ul style="list-style-type: none"> • Viability • Steroid production • Differentiation <p>Oocyte:</p> <ul style="list-style-type: none"> • Spindle analysis • Growth • Nuclear maturation 	oogenesis	
<i>In vitro</i> bovine oocyte maturation assay (bIVM)	Bovine cumulus/oocytes complexes	<p>Monitors the potential adverse effects of chemicals to the maturation process of bovine oocytes. The assay is focused on nuclear configuration changes within the oocyte.</p> <p>Endpoints:</p> <p>Successful achievement of the maturation stage of meiosis up to metaphase II</p>	To identify chemicals which impair the completion of the maturation process of a mammalian oocyte.	Optimisation
Mouse Peri-Implantation Assay (MEPA)	Mouse zygotes	<p><i>In vitro</i> cultures of mouse zygotes are used and the embryo development is observed daily in order to identify any deviations of the timely regulated pre-implantation embryo.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Morphology • Viability • Blastocyst rate (number of blastocysts produced) • Hatching capacity 	<i>In vitro</i> bioassay that allows studying the effect of compounds on the development of the pre-implantation embryo.	Optimisation
<i>In vitro</i> bovine	Bovine	Monitors the effects of chemicals on the formation of	Toxicity assessment	Optimisation

Method	Test system	Description	Applicability	Status
fertilization assay (bIVF)	cumulus/oocytes complexes and cryopreserved bovine sperm.	<p>male and female pronuclei after the penetration of sperm in the oocyte. Oocyte and sperm are exposed to test chemicals.</p> <p>Endpoints:</p> <ul style="list-style-type: none"> • Presence of the sperm head into the oocyte • Presence of two well-formed pronuclei, two polar bodies and the sperm tail. 	during oocyte fertilization process, with possible implications on fertility	

Chapter 10. Table of References with Notes

EU legislation	
REACH Regulation (EC No 1907/2006)	<p>Studies on reproductive and developmental toxicity are required from Annex VIII onwards, i.e. for all substances marketed or manufactured in quantities above 10 t/a. From Annex VIII through X, the standard information requirements are cumulative; requirements at higher tonnage levels add to the information requirements at lower tonnage levels. Standard information requirements include a screening study on reproduction toxicity (OECD TG 421/422) at Annex VIII, a prenatal developmental toxicity study (EU TM B.31 on first species at Annex IX, a two-generation reproductive toxicity study (EU TM B.35) and prenatal developmental toxicity study on an second species at Annex X (if not conducted already at Annex IX). Taken together, at Annex X level, information on reproduction toxicity could be available from a screening reproductive toxicity study, prenatal developmental toxicity studies on two species, two-generation reproductive toxicity study, as well as from repeated dose toxicity studies.</p>
Regulation concerning the marketing and use of biocidal products (EU No 528/2012)	<p>A biocidal product shall not be authorised for use by the general public where it is classified as toxic for reproduction category 1 or 2 (Directive 1999/45/EC); toxic for reproduction category 1A or 1B under Regulation (EC) No 1272/2008; or it has developmental neurotoxic or immunotoxic effects.</p> <p>From the legislation "8.10. Reproductive toxicity ... For evaluation of consumer safety of active substances that may end up in food or feed, it is necessary to conduct toxicity studies by the oral route.</p> <p>If a substance is known to have an adverse effect on fertility, meeting the criteria for classification as Reproductive toxicity Cat 1A or 1B: May damage fertility (H360F), and the available data are adequate to support a robust risk assessment, then no further testing for fertility will be necessary. However, testing for developmental toxicity must be considered if a substance is known to cause developmental toxicity, meeting the criteria for classification as Reproductive toxicity Cat 1A or 1B: May damage the unborn child (H360D), and the available data are adequate to support a robust risk assessment, then no further testing for developmental toxicity will be necessary. However, testing for effects on fertility must be considered.</p>

	<p>8.10.1. Pre-natal developmental toxicity study, preferred species is rabbit; oral route of administration is the preferred route. The study shall be initially performed on one species.</p> <p>8.10.2. Two-generation reproductive toxicity study, rat, oral route of administration is the preferred route. If another reproductive toxicity test is used justification shall be provided. The extended one-generation reproductive toxicity study adopted at OECD level shall be considered as an alternative approach to the multi-generation study.</p> <p>8.10.3. Further pre-natal developmental toxicity study. A decision on the need to perform additional studies on a second species or mechanistic studies should be based on the outcome of the first test (8.10.1) and all other relevant available data (in particular rodent reproductive toxicity studies). Preferred species is rat, oral route of administration</p> <p>For repeated dose toxicity and reproductive toxicity the dose-response relationship shall be assessed for each active substance or substance of concern and, where possible, a NOAEL identified. If it is not possible to identify a NOAEL, the lowest-observed-adverse-effect level (LOAEL) shall be identified. Where appropriate, other dose-effect descriptors may be used as reference values."</p>
Regulation on Plant Protection Products (EC No 1107/2009)	<p>An active substance, safener or synergist shall only be approved if, on the basis of assessment of reproductive toxicity testing carried out in accordance with the data requirements for the active substances, safeners or synergists and other available data and information, including a review of the scientific literature, reviewed by EFSA, it is not or has not to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as toxic for reproduction category 1A or 1B, unless the exposure of humans to that active substance, safener or synergist in a plant protection product, under realistic proposed conditions of use, is negligible, that is, the product is used in closed systems or in other conditions excluding contact with humans and where residues of the active substance, safener or synergist concerned on food and feed do not exceed the default value set in accordance with point (b) of Article 18(1) of Regulation (EC) No 396/2005.</p>

From the Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009:

"5.6 Reproductive Toxicity:

Possible effects on reproductive physiology and the development of progeny shall be investigated and reported concerning the following aspects:

- Impairment of male and female reproductive functions or capacities [...]
- Harmful effects on the progeny, for example any effect interfering with normal development both before and after birth. [...]

Effects accentuated over generations shall be reported.

The active substance and its relevant metabolites shall be measured in milk as a second tier investigation where relevant effects are observed in the offspring or are expected. [...]

Potential neurotoxic, immunotoxic effects and effects potentially related to changes in the hormonal system shall be carefully addressed and reported.

Investigations shall take account of all available and relevant data, including the results of general toxicity studies if relevant parameters [...] are included, as well as knowledge concerning structural analogues to the active substance.

While the standard reference point for treatment responses shall be concurrent control data, historical control data may be helpful in the interpretation of particular reproductive studies. [...]

5.6.1 Generational studies

[...] A reproductive toxicity study in rats over at least two generations shall be reported.

The OECD extended one-generation reproductive toxicity study may be considered as an alternative [...]

Where necessary for a better interpretation of the effects on reproduction and as far as this information is not yet available, supplementary studies may be required to provide information on the affected gender and the possible mechanisms.

	<p>5.6.2. <i>Developmental toxicity studies</i> [...] Developmental toxicity studies shall always be carried out. Developmental toxicity shall be determined for rat and rabbit by the oral route; the rat study shall not be conducted if developmental toxicity has been adequately assessed as part of and extended one-generation reproductive toxicity study. Additional routes may be useful in human risk assessment. [...]</p> <p>When indicated by observations in other studies or the mode of action of the test substance, supplementary studies or information may be required to provide information on the postnatal manifestation of effects such as developmental neurotoxicity."</p>
Cosmetics: Directive 76/768/EEC and Regulation (EC) No 1223/2009	Introduction of the directive resulted in an instant ban on the animal testing of finished consumer products, which in March 2009, was extended to a ban on the animal testing of cosmetic ingredients, irrespective of the availability of alternative methods. Since March 2013, a marketing ban has been in place to reinforce this prohibition.
Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging	Aligns previous EU legislation on classification, packaging and labelling (Dangerous Substance Directive 67/548/EEC) of chemicals with the GHS (Globally Harmonized System of Classification and Labelling of Chemicals). In addition to the differences in the hazard statements that now begin with "H" instead of the old "R" phrases, a very important revolution is in place: submitters are now fully responsible for the classification and assessment of the chemical substances that they either manufacture or import into the EU. CLP still contains in its Annex I a full list of chemicals with mandatory harmonized classification, but it is explicitly required that whenever a new hazard is known, it must be immediately communicated to downstream users and consumers.
US legislation	
Guidelines for Developmental Toxicity Risk Assessment. http://epa.gov/ncct/v-Embryo/files/EPA_Guidelines_Developmental_Toxicity.pdf?p_download_id=4560	The U.S. Environmental Protection Agency (EPA) is today issuing final amended guidelines for assessing the risks for developmental toxicity from exposure to environmental agents. As background information for this guidance, this notice describes the scientific basis for concern about exposure to agents that cause developmental toxicity, outlines the general process for assessing potential risk to humans because of environmental contaminants, summarizes the history of these guidelines, and addresses public and Science Advisory Board comments on the

	<p>1989 “Proposed Amendments to the Guidelines for the Health Assessment of Suspect Developmental Toxicants” [54 FR 9386-9403]. These guidelines, which have been renamed “Guidelines for Developmental Toxicity Risk Assessment” (hereafter “Guidelines”), outline principles and methods for evaluating data from animal and human studies, exposure data, and other information to characterize risk to human development, growth, survival, and function because of exposure prior to conception, prenatally, or to infants and children. These Guidelines amend and replace EPA’s 1986 “Guidelines for the Health Assessment of Suspect Developmental Toxicants” [51 FR 34028-34040] by adding new guidance on the relationship between maternal and developmental toxicity, characterization of the health-related database for developmental toxicity risk assessment, use of the reference dose or reference concentration for developmental toxicity (RfDDT or RfCDT), and use of the benchmark dose approach.</p> <p>In addition, the Guidelines were reorganized to combine hazard identification and dose-response evaluation since these are usually done together in assessing risk for human health effects other than cancer.</p>
<p>Guidelines for Reproductive Toxicity Risk Assessment Published on October 31, 1996, Federal Register 61(212): 56274-56322 http://www.epa.gov/raf/publications/pdfs/REPRO51.PDF</p> <p>(This document replaces the proposed guidelines for Female Reproductive Risk and Proposed guideline for Male Reproductive Risk, both dated June 30, 1988.)</p>	<p>"The U.S. Environmental Protection Agency (EPA) published a document entitled <i>Guidelines for Reproductive Toxicity Risk Assessment</i> (hereafter "Guidelines"). These Guidelines were developed as part of an interoffice guidelines development program by a Technical Panel of the Risk Assessment Forum. They were proposed initially in 1988 as separate guidelines for the female and male reproductive systems. Subsequently, based upon the public comments and Science Advisory Board (SAB) recommendations, changes made included combining those two guidelines, integrating the hazard identification and dose-response sections, assuming as a default that an agent for which sufficient data were available on only one sex may also affect reproductive function in the other sex, expansion of the section on interpretation of female endpoints, and consideration of the benchmark dose approach for quantitative risk assessment. These Guidelines were made available again for public comment and SAB review in 1994. This notice describes the scientific basis for concern about exposure to agents that cause reproductive toxicity, outlines the general process for assessing potential risk to humans from exposure to environmental agents, and addresses Science Advisory Board and public comments on the 1994 <i>Proposed Guidelines for Reproductive Toxicity Risk Assessment</i>. Subsequent reviews have included the Agency’s Risk Assessment Forum and interagency comment by members of</p>

	subcommittees of the Committee on the Environment and Natural Resources of the Office of Science and Technology Policy."
OECD TGs	
OECD TG 414: prenatal development toxicity study for the testing of chemicals (OECD 2001) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788	<p>Studies the effects of prenatal exposure on the pregnant animal and on the developing organism; this may include assessment of maternal effects as well as death, structural abnormalities, or altered growth in the foetus.</p> <p>Period considered: from preimplantation to the day before birth.</p> <ul style="list-style-type: none"> – Endpoints: litter composition (e.g. resorptions, live, dead foetuses), embryonic development, foetal growth, morphological variations and malformations. Functional deficits are not considered. <p>Species: rodent (preferably rat) and non-rodent (preferably rabbit).</p> <p>Standard information requirement for substance under REACH, active substances under Biocide Products Regulations and Regulation on Plant Protection Products.</p>
OECD TG 415: one-generation reproduction toxicity study (OECD 1983) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788	<p>OECD TG 415 is designed to provide general information concerning the effects of a test substance on male and female reproductive performance. Studies the effects on male and female reproductive performance, such as gonadal function, oestrous cycle, mating behaviour, conception, parturition, lactation and weaning. It may also provide preliminary information about developmental toxic effects, such as neonatal morbidity, mortality, behaviour and teratogenesis and to serve as a guide for subsequent tests.</p> <p>Period considered: continuously over one generation.</p> <p>Endpoints: growth, development and viability; pregnancy length and birth outcome; histopathology of sex organs and target organs; and fertility.</p> <p>Preferred species: rat or mouse</p>
OECD TG 416: two-generation reproduction toxicity (OECD 2001) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788	<p>Studies the effects of a substance on the integrity and performance of the male and female reproductive systems, and on the growth and development of the offspring, including gonadal function, the oestrus cycle, mating behaviour, conception, gestation, parturition, lactation, and weaning, and the growth and development of the offspring. It may provide information on neonatal morbidity, mortality, and preliminary data on prenatal and postnatal developmental toxicity.</p> <p>Period considered: continuously over two or several generations.</p> <p>Endpoints: growth, development and viability; pregnancy length and birth outcome; histopathology of sex organs and target organs; fertility; and oestrus cyclicity and sperm quality.</p>

	<p>Preferred species: the rat. Standard information requirement for substances under REACH, active substances under Biocide Products Regulation, and Regulation on Plant Protection Products.</p>
<p>OECD TG 421: reproduction/developmental toxicity screening test (OECD 1995) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>OECD TG 421 can be used to provide <u>initial information</u> on possible effects on reproduction and/or development. This test does not provide complete information on all aspects of reproduction and development. Generates preliminary information concerning the effects of a substance on male and female reproductive performance such as gonadal function, mating behaviour, conception, development of the conceptus and parturition. It is not an alternative to, nor does it replace the Test Guidelines 414, 415 and 416. Positive results are useful for initial hazard assessment and contribute to decisions with respect to the necessity and timing of additional testing. Period: from 2 weeks prior to mating until day 4 postnatally. Endpoints: fertility; pregnancy length and birth outcome; histopathology of sex organs and target organs; foetal and pup growth and survival until day 3. Preferred species: the rat. Standard information requirement under the REACH regulation.</p>
<p>OECD TG 422: combined repeated dose toxicity study with the reproduction/developmental toxicity screening test (OECD 1996) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>OECD TG 422 provides information on the possible health hazards likely to arise from repeated exposure over a relatively limited period of time. It also can be used to provide initial information on possible effects on male and female reproductive performance. This test does not provide complete information on all aspects of reproduction and development. Apart from gonadal function, mating behaviour, conception, development of the conceptus and parturition, the Guideline also places emphasis on neurological effects. Useful as part of the initial screening of chemicals for which little or no toxicological information is available and can serve as an alternative to conducting two separate tests for repeated dose toxicity (TG 407) and reproduction/developmental toxicity (TG 421), respectively. It can also be used as a dose range finding study for more extensive reproduction/developmental studies, or when otherwise considered relevant. It will not provide evidence for definite claims of no reproduction/developmental effects. Period: from 2 weeks prior to mating until day 4 postnatally. Endpoints: fertility; pregnancy length and birth outcome; histopathology of sex organs and target organs and brain; foetal and pup growth and survival until day</p>

	<p>3. Preferred species: the rat. Alternative to standard requirement for 28 day repeated dose toxicity study and OECD TG 421 under the REACH regulation.</p>
<p>OECD test guideline 426: developmental neurotoxicity study (OECD 2007) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>Study the potential functional and morphological effects on the developing nervous system of the offspring of repeated exposure to a substance during in utero and early postnatal development. It can be conducted as a separate study, incorporated into a reproductive toxicity and/or adult neurotoxicity study (e.g. TG 415, 416, 424), or added onto a prenatal developmental toxicity study (e.g. TG 414). Period: during pregnancy and lactation. Endpoints: pregnancy length and birth outcome; physical and functional maturation; behavioural changes due to CNS and PNS effects; and brain weights and neuropathology. Preferred species: the rat. It is only regulatory required for the evaluation of agrochemicals and as an additional data set in Biocide Products Regulation. .</p>
<p>OECD test guideline 440: uterotrophic bioassay in rodents: a short-term screening test for oestrogenic properties (OECD 2007) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>This <i>in vivo</i> test in immature or ovariectomised animals evaluates the ability of a chemical to elicit biological endocrine disruption activities consistent with agonists or antagonists of natural oestrogens (e.g. 17β-estradiol). It is based on the increase in uterine weight or uterotrophic response. The uterus responds to oestrogens with an increase in weight due to water imbibition, followed by a weight gain due to tissue growth. Endpoint: uterotrophic response to oestrogens. Preferred species: rat (or mature mice).</p>
<p>OECD test guideline 441: Hershberger bioassay in rats: a short-term screening assay for (anti-) androgenic properties (OECD 2009) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>This <i>in vivo</i> test in non-intact animals evaluates the ability of a chemical to elicit biological endocrine disruption activities consistent with androgen agonists, antagonists or 5 α-reductase inhibitors. Endpoints: changes in weight of five androgen-dependent tissues in the castrate-peripubertal male rat Preferred species: the rat.</p>
<p>OECD test guideline 443: extended one-generation reproductive toxicity study (OECD 2011) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>This study has a flexible modular study design with cohorts for reproductive/developmental toxicity (Cohorts 1A and 1B), developmental neurotoxicity (Cohorts 2A and 2B), and developmental immunotoxicity (Cohort 3). Cohort 1B may be extended to include mating of the F1 animals and</p>

	<p>production of F2 generation. The study focuses on effects on F1 animals and it is optimised to assess effects related to certain endocrine modes of action.</p> <p>Shall be considered as an alternative approach to multi-generation study in Biocide Products Regulation.</p>
<p>OECD test guideline 455: the stably transfected human oestrogen receptor-α transcriptional activation assay for detection of oestrogenic agonist-activity of chemicals (OECD 2009c) http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788</p>	<p>This <i>in vitro</i> assay evaluates the transcriptional activation mediated by the hERα of oestrogen responsive genes, a process considered to be one of the key mechanisms of possible endocrine disruption related health hazards. The assay provides mechanistic information and can be used for screening and prioritisation purposes of oestrogenic compounds.</p> <p>Endpoint: induction of hERα-mediated transactivation of luciferase gene expression.</p> <p>Test system: the hERα-HeLa-9903 cell line derived from a human cervical tumour and stably transfected.</p> <p>It is only used in some cases for cosmetic testing.</p>
Mechanistic understanding	
<p>Perkins EJ, Ankley GT, Crofton KM, Garcia-Reyero N, Lalone CA, Johnson MS, Tietge JE & Villeneuve DL (2013). Current perspectives on the use of alternative species in human health and ecological hazard assessments. <i>Environmental Health Perspectives</i> 121: 1002–1010</p>	<p>The authors reviewed and compared data spanning from high throughput <i>in vitro</i> assays to fish reproductive tests for seven chemicals and further investigated whether human-focused assays can be predictive of chemical hazards in the environment. By using several examples, the authors demonstrate that pathway-based analysis of chemical effects provides new opportunities to use alternative models (non-mammalian species, <i>in vitro</i> tests) to support decision making while reducing animal use and associated costs. The authors postulate that, using a pathway-based hazard assessment approach, data from multiple species and non-animal alternative models are equally valuable for both ecological and human health hazard assessment. Likewise, the paper describes how alternative models can be predictive of effects of human health concern (e.g., endocrine disruption) and link chemicals to toxicity pathways, or modes of action, in both mammals and ecological species. Finally, the way how dose-dependent effects in alternative models can be translated using a pathway-based measure to chemical hazard levels that are similar to those generated using mammalian species in chronic tests is documented. By these examples the authors highlight the scientifically credible foundation that supports the predictive application and/or extrapolation of pathway-based toxicological data across species.</p>
<p>Van Dartel, D. a M., & Piersma, A. H. (2011). The embryonic stem cell test combined with toxicogenomics as an alternative testing model</p>	<p>In the present review, the progress made with regard to the prediction of developmental toxicity using the EST combined with transcriptomics is</p>

<p>for the assessment of developmental toxicity. <i>Reproductive toxicology</i>, 32(2), 235–44.</p>	<p>presented. Furthermore, the developments of additional aspects required for further optimization of the EST, including kinetics, the use of human embryonic stem cells (ESC) and computational toxicology are discussed. Finally, the current and future use of the EST model for prediction of developmental toxicity in testing strategies and in regulatory toxicity evaluations is discussed.</p>
<p>Yozzo, K. L., McGee, S. P., & Volz, D. C. (2013). Adverse outcome pathways during zebrafish embryogenesis: a case study with paraoxon. <i>Aquatic toxicology</i>, 126, 346–54.</p>	<p>The objective of this study was to develop an adverse outcome pathway (AOP) that provided quantitative linkages across levels of biological organization during zebrafish embryogenesis. Using a well-studied reference chemical, this study highlights the potential challenges in developing quantitative AOPs to support chemical screening and prioritization strategies.</p>
<p>De Esch, C., Sliker, R., Wolterbeek, A., Woutersen, R., & de Groot D. (2012). Zebrafish as a potential model for developmental neurotoxicity testing: A mini review. <i>Neurotoxicology and Teratology</i>, 34: 545-553.</p>	<p>The zebrafish is a powerful toxicity model; biochemical assays can be combined with observations at a structural and functional level within one individual. This mini review summarises the potency of zebrafish as a model for developmental neurotoxicity screening, and its possibilities to investigate working mechanisms of toxicants. The use of zebrafish in toxicity research can ultimately lead to the refinement or reduction of animal use.</p>
<p><i>In vitro</i> test methods</p>	
<p>Louisse, J., Verwei, M., Woutersen A. R., Blaauboer J. B., Rietjens MCM. I. (2012). Toward <i>in vitro</i> biomarkers for developmental toxicity and their extrapolation to the <i>in vivo</i> situation. <i>Expert Opin Drug Metab Toxicol</i>, 8(1), 11-27</p>	<p>The present review gives an overview of alternative assays, as described in the literature, for <i>in vivo</i> developmental toxicity, including the effects (readouts) assessed in these assays. The authors discuss how these data may be used to obtain relevant biomarkers for <i>in vivo</i> developmental toxicity, and how <i>in vitro</i> effect data can be translated to the <i>in vivo</i> situation using physiologically based kinetic (PBK) modelling.</p> <p>Relevance of readouts in <i>in vitro</i> developmental toxicity assays as predictive biomarkers for <i>in vivo</i> developmental toxicity should be evaluated by comparing the obtained <i>in vitro</i> effect concentrations with <i>in vivo</i> internal concentrations at dose levels causing developmental toxicity. Extrapolation of the <i>in vitro</i> effect concentrations to <i>in vivo</i> dose levels using PBK modelling (i.e., reverse dosimetry) is promising in its use to derive points of departure for risk assessment, enabling the use of <i>in vitro</i> toxicity data in the safety assessment of compounds.</p>

<p>Adler, S., Basketter, D., Creton, S., Pelkonen, O., Van Benthem, J., Zuang, V., et al. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. <i>Archives of Toxicology</i> 85: 367–485.</p>	<p>This report provides the findings of a panel of scientific experts tasked with assessing the availability of alternative methods to animal testing in view of the full marketing ban for cosmetic products and ingredients tested on animals in Europe.</p>
<p>Theunissen, Peter T, Robinson, J. F., Pennings, J. L. a, De Jong, E., Claessen, S. M. H., Kleinjans, J. C. S., & Piersma, A. H. (2012). Transcriptomic concentration-response evaluation of valproic acid, cyproconazole, and hexaconazole in the neural embryonic stem cell test (ESTn). <i>Toxicological sciences: an official journal of the Society of Toxicology</i> 125: 430–438.</p>	<p>The concentration-dependent effects of three known (neuro) developmental toxicants, two triazoles, cyproconazole (CYP) and hexaconazole (HEX), and the anticonvulsant valproic acid (VPA) were investigated using a transcriptomic approach. Compound effects on gene expression during neural differentiation and corresponding regulated gene ontology (GO) terms were identified after 24 h of exposure in relation to morphological changes on day 11 of culture. Concentration-dependent responses on individual gene expression and on biological processes were determined for each compound, providing information on mechanism and concentration-response characteristics. All compounds caused enrichment of the embryonic development process. CYP and VPA but not HEX significantly enriched the neuron development process. Furthermore, specific responses for triazole compounds and VPA were observed within the GO-term sterol metabolic process. The incorporation of transcriptomics in the ESTn was shown to enable detection of effects, which precede morphological changes and provide a more sensitive measure of concentration-dependent effects as compared with classical morphological assessments. Furthermore, mechanistic insight can be instrumental in the extrapolation of effects in the ESTn to human hazard assessment.</p>
<p>Zhang, C., Cao, J., Kenyon, J. R., Panzica-kelly, J. M., Gong, L., & Augustine-rauch, K. (2012). Development of a Streamlined Rat Whole Embryo Culture Assay for Classifying Teratogenic Potential of Pharmaceutical Compounds, 127(2), 535–546.</p>	<p>In this paper a novel rat whole embryo culture (rWEC) teratogenicity assay applying a simplified experimental design and statistical prediction model is presented. The described assay holds promise for reduced animal requirements and increased throughput with low prediction error rate for classifying teratogenic potential of compounds. A total of 70 compounds (38 teratogens and 32 non-teratogens) were evaluated, and the prediction model was generated from a dataset of 59 compounds. The rWEC assay requires only one test concentration and three structural endpoints, which are used in a recursive partition model for classifying teratogenic liability. The model fitting concordance between the WEC assay and in vivo outcome was 83%. The predictivity for future compounds was evaluated by using two statistical methods. The overall estimate for prediction concordance is 74% (SD 5.2%). There is no statistically significant difference (p value > 0.50) in the predictivity between this model and the model supporting</p>

	<p>ECVAM validated WEC assay with predictivity of 80% (SD 10.6%). Overall, the streamlined WEC assay is estimated to reduce animal use and operational costs by more than 50% and it substantially improves results turnaround with no loss of predictivity.</p>
<p>Uysal, H., Semerdoken, S., Colak, D. A., & Ayar, A. (2013). The hazardous effects of the three natural food dyes on developmental stages and longevity of <i>Drosophila melanogaster</i>. <i>Toxicology and Industrial Health</i>. Epub ahead of print, 1 March 2013. doi: 10.1177/0748233713480206.</p>	<p>In this study, the toxic effects of three different natural food dyes (carmine, turmeric and annatto) on 72 ± 4 h larvae of <i>Drosophila melanogaster</i> were investigated. For this purpose, four different application doses were chosen by means of preliminary studies. It was determined that larval mortality increased with increasing concentration in the application groups and the toxicity order was carmine > turmeric > annatto. Moreover, it was observed that the survival rate was highest in the control with 98% and lowest in maximal used dose of carmine with 16%. In addition, the average lifespan of the adult individuals obtained from third instar larvae was also studied. According to the obtained results, when both the developmental period from larvae into adults and the lifespan of the developing adults were compared with the control group, the food dyes were found to be toxic and the toxicity order of carmine > turmeric > annatto was identified.</p>
<p>Goussen, B., Parisot, F., Beaudouin, R., Dutilleul, M., Buisset-Goussen, A., Péry, A. R. R., & Bonzom, J.-M. (2013). Consequences of a multi-generation exposure to uranium on <i>Caenorhabditis elegans</i> life parameters and sensitivity. <i>Ecotoxicology (London, England)</i>.</p>	<p>In this paper a 16-generation study was performed to assess the uranium exposure effects of two populations of the ubiquitous nematode <i>Caenorhabditis elegans</i>. Several generations were selected to assess growth, reproduction, survival, and dose-responses relationships, through exposure to a range of concentrations with all endpoints measured daily. The experiment showed an adaptation of individuals to experimental conditions (increase of maximal length and decrease of fecundity) for both populations. Moreover an increase of adverse effects (reduction of growth and fertility) as a function of uranium concentration was observed. The group also reported the emergence of population differentiation for reproduction traits. In contrast, no differentiation was observed on growth traits. These results confirm the importance of assessing environmental risk related to pollutant through multi-generational studies.</p>
<p>Zimmer, B., Lee, G., Balmer, N. V, Meganathan, K., Sachinidis, A., Studer, L., & Leist, M. (2012). Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. <i>Environmental health perspectives</i>, 120(8), 1116–22.</p>	<p>This paper describes a robust and widely applicable human-relevant neural crest (NC) function assay that would allow for sensitive screening of environmental toxicants and defining toxicity pathways. The NC cells from human embryonic stem cells were generated, and a migration assay of NC cells (MINC assay) for testing environmental toxicants was established. Methylmercury, valproic acid, and lead-acetate affected the migration of NC cells more potently than migration</p>

	<p>of other cell types. The MINC assay correctly identified the NC toxicants triadimefon and triadimenol. Additionally, it showed different sensitivities to various organic and inorganic mercury compounds. Using the MINC assay and applying classic pharmacologic inhibitors and large-scale microarray gene expression profiling, the authors found several signaling pathways that are relevant for the migration of NC cells. The MINC assay faithfully models human NC cell migration, and it reveals impairment of this function by developmental toxicants with good sensitivity and specificity.</p>
<p>Li, H., Van Ravenzwaay, B., Rietjens, I. M. C. M., & Louise, J. (2013). Assessment of an in vitro transport model using BeWo b30 cells to predict placental transfer of compounds. <i>Archives of toxicology</i>. Archives of Toxicology 87(9):1661-1669.</p>	<p>The human <i>ex vivo</i> placental perfusion model has regularly been used to study the transplacental transport of compounds. However, this method is laborious and dependent on the presence of fresh human placenta, hampering its use for the assessment of large numbers of compounds. An <i>in vitro</i> model for the placental barrier using BeWo b30 cells may provide an alternative to the <i>ex vivo</i> system. The present study aims to assess whether such an <i>in vitro</i> model could be used to reliably predict placental transfer. To this end, BeWo b30 cells, derived from a human choriocarcinoma, were grown on transwell insert to form a cell layer, separating an apical maternal compartment from a basolateral fetal compartment. For a set of nine selected model compounds, including the reference compound antipyrine, the transport velocity from the apical to the basolateral compartment was determined. Relative transport rates obtained were compared with the transfer indices (a measure for the transport relative to antipyrine) of these compounds obtained in <i>ex vivo</i> placental perfusion studies as reported in the literature. The relative transport rates in the <i>in vitro</i> BeWo model were in good correlation ($R^2 = 0.95$) with the transfer indices reported for the <i>ex vivo</i> model. This indicates that the BeWo model could be a valuable <i>in vitro</i> model for prediction of placental transfer of compounds.</p>
<p>Myllynen P., Immonen E., Kummu M., Vahakangas K. (2009) Developmental expression of drug metabolizing enzymes and transporter proteins in human placenta and fetal tissues. <i>Expert Opin Drug Metab Toxicol</i>. 5,1483–1499</p>	<p>Over the decades several <i>ex vivo</i> and <i>in vitro</i> models which utilize delivered human placenta have been developed to study various placental functions. The use of models originating from human placenta to study transplacental transfer and related mechanisms is an attractive option because human placenta is relatively easily available for experimental studies. After delivery, placenta has served its purpose and is usually disposed of. The purpose of this review is to give an overview of the use of human placental models for the studies on human placental transfer and related mechanisms such as transporter functions and xenobiotic metabolism. Human placental perfusion, the most commonly used</p>

	continuous cell lines, primary cells and tissue culture, as well as subcellular fractions are briefly introduced and their major advantages and disadvantages are discussed.
Gustafson, AL, Stedman DB, Ball J, Hillegass JM, Flood A, Zhang CX, Panzica-Kelly J, Cao J, Coburn A, Enright BP, Tornesi MB, Hetheridge M & Augustine-Rauch KA (2012). Inter-Laboratory Assessment of a Harmonized Zebrafish Developmental Toxicology Assay – Progress Report on Phase I. <i>Reproductive Toxicology</i> 33: 155-164.	This report provides a progress update of a consortium effort to develop a harmonized zebrafish developmental toxicity assay. Twenty non-proprietary compounds (10 animal teratogens and 10 animal non-teratogens) were evaluated blinded in 4 laboratories. Zebrafish embryos from pond-derived and cultivated strain wild types were exposed to the test compounds for 5 days and subsequently evaluated for lethality and morphological changes. Each of the testing laboratories achieved similar overall concordance to the animal data (60–70%). Subsequent optimization procedures to improve the overall concordance focused on compound formulation and test concentration adjustments, chorion permeation and number of replicates. These optimized procedures were integrated into a revised protocol and all compounds were retested in one lab using embryos from pond-derived zebrafish and achieved 85% total concordance. To further assess assay performance, a study of additional compounds is currently in progress at two laboratories using embryos from pond-derived and cultivated-strain wild type zebrafish.
Selderslaghs IWT, Blust R & Witters H (2012). Feasibility study of the zebrafish assay as an alternative method to screen for developmental toxicity and embryotoxicity using a training set of 27 compounds. <i>Reproductive Toxicology</i> 33:142-154.	The zebrafish embryo/larva was studied as an alternative for animal testing for developmental toxicity and embryotoxicity and a training set of 27 compounds was evaluated with a standardized protocol. The classification of compounds in the zebrafish embryo/larva assay, based on a prediction model using a TI (teratogenic index) cut-off value of 2, was compared to available animal and human data. When comparing the classification of compounds in the zebrafish embryo/larva assay to available animal classification, a sensitivity of 72% and specificity of 100% were obtained. The predictive values obtained in comparison to a limited set of human data were 50, 60% respectively for teratogens, non-teratogens. Overall, it was demonstrated that the zebrafish embryo/larva assay, may be used as a screening tool for prioritization of compounds and could contribute to reduction of animal experiments in the field of developmental toxicology.
In silico models	
Arena VC, Sussman NB, Mazumdar S, Yu S & Macina OT (2004). The utility of Structure-activity relationship (SAR) models for prediction and covariate selection in developmental toxicity:	Structure-activity relationship (SAR) models can be used to predict the biological activity of potential developmental toxicants whose adverse effects include death, structural abnormalities, altered growth and functional deficiencies in the

<p>comparative analysis of logistic regression and decision tree models. SAR and QSAR in Environmental Research 15: 1-18.</p>	<p>developing organism. Physicochemical descriptors of spatial, electronic and lipophilic properties were used to derive SAR models by two modelling approaches, logistic regression and Classification and Regression Tree (CART), using a new developmental database of 293 chemicals (FDA/TERIS). Both single models and ensembles of models (termed bagging) were derived to predict toxicity. Assessment of the empirical distributions of the prediction measures was performed by repeated random partitioning of the data set. Results showed that both the decision tree and logistic regression derived developmental SAR models exhibited modest prediction accuracy. Bagging tended to enhance the prediction accuracy and reduced the variability of prediction measures compared to the single model for CART-based models but not consistently for logistic-based models. Prediction accuracy of single logistic-based models was higher than single CART-based models but bagged CART-based models were more predictive. Descriptor selection in SAR for the understanding of the developmental mechanism was highly dependent on the modelling approach. Although prediction accuracy was similar in the two modelling approaches, there was inconsistency in the model descriptors.</p>
<p>Sussman, N. B., Arena, V. C., Yu, S., Mazumdar, S., and Thampatty, B. P. 2003. Decision Tree SAR Models for Developmental Toxicity Based on an FDA/TERIS Database. <i>SAR and QSAR in Environmental Research</i> 14 (2):83-96.</p>	<p>Structure-activity relationships (SARs) are models that could be used to efficiently predict the biological activity of potential developmental toxicants. However, at this time, no adequate SAR models of developmental toxicity are available for risk assessment. In the present study, a new developmental database was compiled by combining toxicity information from the Teratogen Information System (TERIS) and the Food and Drug Administration (FDA) guidelines. The authors implemented a decision tree modelling procedure, using Classification and Regression Tree (CART) software and a model ensemble approach termed bagging. The authors then assessed the empirical distributions of the prediction accuracy measures of the single and ensemble-based models, achieved by repeating our modelling experiment many times by repeated random partitioning of the working database. The decision tree developmental SAR models exhibited modest prediction accuracy. Bagging tended to enhance the accuracy of prediction. Also, the model ensemble approach reduced the variability of prediction measures compared to the single model approach.</p>
<p>Novic M & Vracko M (2010). QSAR Models for Reproductive Toxicity and Endocrine Disruption Activity. <i>Molecules</i> 2010, 15.</p>	<p>This review describes some QSAR modelling approaches for reproductive toxicity. In the first example the authors describe the CAESAR model for prediction of reproductive toxicity; the second example shows a classification</p>

	<p>model for endocrine disruption potential based on counter propagation artificial neural networks; the third example shows a modelling of relative binding affinity to rat oestrogen receptor, and the fourth one shows a receptor dependent modelling experiment.</p>
<p>Lo Piparo E & Worth A (2010). Review of QSAR Models and Software Tools for predicting Developmental and Reproductive Toxicity. JRC report EUR 24522 EN. Publications Office of the European Union. Available from: http://publications.jrc.ec.europa.eu/repository</p>	<p>This JRC report provides a state-of-the-art review of available computational models for developmental and reproductive toxicity, including Quantitative Structure-Activity Relationship (QSARs) and related estimation methods such as decision tree approaches and expert systems. At present, there are relatively few models for reproductive toxicity endpoints (development and fertility), and those available have limited applicability domains. This situation is partly due to the biological complexity of the endpoint, which covers many incompletely understood mechanisms of action, and partly due to the paucity and heterogeneity of high quality data suitable for model development. In contrast, there is an extensive and growing range of software and literature models for predicting endocrine-related activities, in particular models for oestrogen and androgen activity. There is a considerable need to further develop and characterise <i>in silico</i> models for reproductive toxicity, and to explore their applicability in a regulatory setting.</p>
<p>Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D, Benfenati E (2010). CAESAR models for developmental toxicity. Chemistry Central Journal 4, Suppl 1:S4.</p>	<p>The REACH legislation requires assessment of a large number of chemicals in the European market for several endpoints. Developmental toxicity is one of the most difficult endpoints to assess, on account of the complexity, length and costs of experiments. Following the encouragement of QSAR (<i>in silico</i>) methods provided in REACH itself, the CAESAR project has developed several models.</p> <p>Two QSAR models for developmental toxicity have been developed, using different statistical/mathematical methods. Both models performed well. The first makes a classification based on a random forest algorithm, while the second is based on an adaptive fuzzy partition algorithm. The first model has been implemented and inserted into the CAESAR on-line application, which is java-based software that allows everyone to freely use the model.</p> <p>The CAESAR QSAR models have been developed with the aim to minimize false negatives in order to make them more usable for REACH. The CAESAR on-line application ensures that both industry and regulators can easily access and use the developmental toxicity model (as well as the models for the other four</p>

<p>Cronin MTD & Worth AP (2008). (Q)SARs for Predicting Effects Relating to Reproductive Toxicity. <i>QSAR & Combinatorial Science</i> 27: 91-100.</p>	<p>endpoints).</p> <p>This paper reviews the current status of structure-based methods for predicting adverse reproductive effects in mammals. The methods described include QSARs, expert systems and the less formalised approaches of read-across and chemical categories. There are a number of problems with applying QSARs to reproductive toxicology notably the complexity, subtlety and sometimes ill-defined nature of the endpoint and lack of data available for modelling. A small number of QSARs have been developed for individual classes of compounds on well-defined effects. These are supplemented by expert systems approaches (e.g. DEREK for Windows, TOPKAT, MultiCASE, PASS, OECD QSAR Application Toolbox) for a variety of endpoints associated with reproductive toxicology. By far the largest, and best developed, group of models are those for receptor binding effects related to endocrine disruption, in particular to the oestrogen receptor and, to a lesser extent, the androgen receptor. Strategies to improve predictive capabilities for reproductive toxicology are also suggested.</p>
<p>Worth A, Fuart-Gatnik M, Lapenna S & Serafimova R (2011). Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Report produced for EFSA. Available from: http://www.efsa.europa.eu/en/supporting/pub/169e.htm</p>	<p>This is the final report of a project carried out by the European Commission's Joint Research and sponsored by the European Food Safety Authority with the overall aim of evaluating the potential applicability of computational methods for predicting adverse developmental and neurotoxic effects in the dietary risk assessment of pesticides.</p> <p>While the toxicological profile of the parent active substance is fully characterised through the experimental studies required by EU legislation, only very limited toxicological data are usually available for their metabolites and degradates. For reasons of efficiency and animal welfare, computational methods based on structure-activity analysis and read-across are being investigated for their applicability in assessing the toxicological relevance of metabolites and degradates of pesticide active substances. The ability to reliably predict the presence and absence of short-term effects of concern, and in particular developmental toxicity and neurotoxicity, would have a positive impact on the way pesticide risk assessments are currently carried out by reducing the need for toxicity testing on metabolites and degradates as well as the need to conduct short-term exposure assessments.</p> <p>In this study, the ability of selected Quantitative Structure-Activity Relationship (QSAR) tools to predict developmental and neurotoxicity was analysed, and a stepwise approach based on the use of QSAR analysis and read-across was</p>

	<p>proposed as a possible way of supporting, alongside other non-testing approaches such as the Threshold of Toxicological Concern (TTC) approach, the assessment of pesticide metabolites and degradates in terms of their toxicological relevance. In this stepwise approach, QSAR tools are used in a preliminary step to identify toxic chemicals, while read-across is applied, in cases where a chemical is predicted by QSAR to be non-toxic, as a means of distinguishing between true and false negatives. This approach is shown to improve the overall ability to distinguish between toxic and non-toxic chemicals compared with the use of individual tools.</p>
<p>Martin MT, Knudsen TB, Reif DM, Houck KA, Judson RS, Kavlock RJ, Dix DJ (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. <i>Biol Reprod.</i> 85(2):327-339.</p>	<p>The U.S. Environmental Protection Agency's ToxCast research program uses high throughput screening (HTS) for profiling bioactivity and predicting the toxicity of large numbers of chemicals. ToxCast Phase I tested 309 well-characterized chemicals in more than 500 assays for a wide range of molecular targets and cellular responses. Of the 309 environmental chemicals in Phase I, 256 were linked to high-quality rat multigeneration reproductive toxicity studies in the relational Toxicity Reference Database. Reproductive toxicants were defined here as having achieved a reproductive lowest-observed-adverse-effect level of less than 500 mg/kg/day. Eighty-six chemicals were identified as reproductive toxicants in the rat, and 68 of those had sufficient <i>in vitro</i> bioactivity to model. Each assay was assessed for univariate association with the identified reproductive toxicants. Significantly associated assays were linked to gene sets and used for the subsequent predictive modelling. Using linear discriminant analysis and fivefold cross-validation, a robust and stable predictive model was produced capable of identifying rodent reproductive toxicants with 77% ± 2% and 74% ± 5% (mean ± SEM) training and test cross-validation balanced accuracies, respectively. With a 21-chemical external validation set, the model was 76% accurate, further indicating the model's potential for prioritizing the many thousands of environmental chemicals with little to no hazard information. The biological features of the model include steroidal and nonsteroidal nuclear receptors, cytochrome P450 enzyme inhibition, G protein-coupled receptors, and cell signalling pathway readouts-mechanistic information suggesting additional targeted, integrated testing strategies and potential applications of <i>in vitro</i> HTS to risk assessment.</p>
<p>Knudsen TB, Kleinstreuer NC (2011). Disruption of embryonic vascular development in predictive toxicology. <i>Birth Defects Res C</i></p>	<p>Toxicity testing in the 21st century is moving toward using high-throughput screening assays to rapidly test thousands of chemicals against hundreds of</p>

<p>Embryo Today 93(4):312-323.</p>	<p>molecular targets and biological pathways, and to provide mechanistic information on chemical effects in human cells and small model organisms. First-generation predictive models for prenatal developmental toxicity have revealed a complex web of biological processes with many connections to vasculogenesis and angiogenesis. This review examines disruption of embryonic vascular development as a potential adverse outcome pathway leading to developmental toxicity. The authors briefly review embryonic vascular development and important signals for vascular development (local growth factors and cytokines such as vascular endothelial growth factor-A and TGF-beta, components in the plasminogen activator system, and chemotactic chemokines). Genetic studies have shown that perturbing these signals can lead to varying degrees of adverse consequences, ranging from congenital angiodyplasia to foetal malformations and embryoletality. The molecular targets and cellular behaviours required for vascular development, stabilization and remodeling are amenable to <i>in vitro</i> evaluation. Evidence for chemical disruption of these processes is available for thalidomide, estrogens, endothelins, dioxin, retinoids, cigarette smoke, and metals among other compounds. Although not all compounds with developmental toxicity show an <i>in vitro</i> vascular bioactivity signature, many 'putative vascular disruptor compounds' invoke adverse developmental consequences. As such, an adverse outcome pathway perspective of embryonic vascular development can help identify useful information for assessing adverse outcomes relevant to risk assessment and efficient use of resources for validation.</p>
<p>Martin MT, Knudsen TB, Judson RS, Kavlock RJ, Dix DJ (2012). Economic benefits of using adaptive predictive models of reproductive toxicity in the context of a tiered testing program. Syst Biol Reprod Med. 58(1): 3-9.</p>	<p>A predictive model of reproductive toxicity, as observed in rat multigeneration reproductive (MGR) studies, was previously developed using high throughput screening (HTS) data from 36 <i>in vitro</i> assays mapped to 8 genes or gene-sets from Phase I of USEPA ToxCast research program, the proof-of-concept phase in which 309 toxicologically well characterized chemicals were tested in over 500 HTS assays. The model predicted the effects on male and female reproductive function with a balanced accuracy of 80%. In a theoretical examination of the potential impact of the model, two case studies were derived representing different tiered testing scenarios to: 1) screen-out chemicals with low predicted probability of effect; and 2) screen-in chemicals with a high probability of causing adverse reproductive effects. 'Testing cost efficiency' was defined as the total cost divided by the number of positive chemicals expected in the definitive guideline toxicity study. This would approach \$2.11 M under the current practice.</p>

	<p>Under case study 1, 22% of the chemicals were screened-out due to low predicted probability of adverse reproductive effect and a misclassification rate of 12%, yielding a test cost efficiency of \$1.87 M. Under case study 2, 13% of chemicals were screened-in yielding a testing cost efficiency of \$1.13 M per test-positive chemical. Applying the model would also double the total number of positives identified. It should be noted that the intention of the case studies is not to provide a definitive mechanism for screening-in or screening-out chemicals or account for the indirect costs of misclassification. The case studies demonstrate the customizability of the model as a tool in chemical testing decision-making. The predictive model of reproductive toxicity will continue to evolve as new assays become available to fill recognized biological gaps and will be combined with other predictive models, particularly models of developmental toxicity, to form an initial tier to an overarching integrated testing strategy.</p>
<p>Testing Strategies</p>	
<p>Grindon C, Combes R, Cronin MT, Roberts DW, Garrod JF (2008). An integrated decision-tree testing strategy for developmental and reproductive toxicity with respect to the requirements of the EU REACH legislation. <i>ATLA</i> 36 (1):123–138</p>	<p>This paper presents some results of a joint research project, sponsored by Defra and conducted by FRAME and Liverpool John Moores University, on the status of alternatives to animal testing with regard to the European Union REACH system for the safety testing and risk assessment of chemicals. The project covered all the main toxicity endpoints associated with REACH.</p> <p>This paper focuses on the prospects for the use of alternative methods (both <i>in vitro</i> and <i>in silico</i>) in reproductive toxicity testing (both fertility and development). It considers many tests based on primary cells and cell lines, and the available expert systems and QSARs for developmental and reproductive toxicity, and also covers tests for endocrine disruption. Ways in which reduction and refinement measures can be used are also discussed by the authors. Decision-tree style integrated testing strategies are also proposed for developmental and reproductive toxicity and for endocrine disruption, followed by a number of recommendations for the future facilitation of reproductive toxicity testing, with respect to human risk assessment.</p>
<p>Robinson, J. F., & Piersma, A. H. (2013). Toxicogenomic approaches in developmental toxicology testing. <i>Methods in molecular biology</i>, 947, 451–73.</p>	<p>This review presents the basic framework of conducting toxicogenomic investigations in the field of developmental toxicology, providing examples of biological and technical factors that may influence response and interpretation. Furthermore, current diverse applications of toxicogenomic-based approaches in developmental toxicity testing, including exposure-response characterization</p>

	<p>(dose and duration), chemical classification studies, and cross-model comparisons study designs are analysed. The intention of this review is to guide scientists through the challenging and complex structure of conducting toxicogenomic analyses, while considering the many applications of using toxicogenomics in study designs and the future of these types of "omics" approaches in developmental toxicology. The emergence of toxicogenomic applications provides new tools to characterize and potentially predict potency of causing malformations. However, due to the vast number of experimental and statistical procedural steps, toxicogenomic studies are challenging.</p>
<p>Theunissen, P T, Schulpen, S. H. W., Van Dartel, D. A. M., Hermsen, S. A. B., Van Schooten, F. J., & Piersma, A. H. (2010). An abbreviated protocol for multilineage neural differentiation of murine embryonic stem cells and its perturbation by methyl mercury. <i>Reproductive Toxicology</i> 29(4), 383–92.</p>	<p>This paper presents a test system for assessing neurodevelopmental toxicity using differentiating embryonic stem cells. The authors claim to have advanced previously established methods by merging, modifying and abbreviating the original 20-day protocol into a more efficient 13-day neural differentiation protocol. Using morphological observation, immunocytochemistry, gene expression and flow cytometry, it was shown that predominantly multiple lineages of neuroectodermal cells were formed in their protocol and to a lower extent, endodermal and mesodermal differentiated cell types. This abbreviated protocol has the potential to be used as an advanced screening method using morphology in combination with selected differentiation markers aimed at predicting neurodevelopmental toxicity. Finally, using the modified protocol differential sensitivity to a model developmental neurotoxicant, methyl mercury, was shown.</p>
<p>Rovida, C., Longo, F., & Rabbit, R. R. (2011). t 4 Report * How are Reproductive Toxicity and Developmental Toxicity Addressed in REACH Dossiers? <i>Altex</i> 28: 273–294.</p>	<p>In the present paper, data from 400 submitted dossiers were analyzed to check for compliance with REACH Regulations and published guidelines. The dossiers were randomly selected among full phase-in registered substances and the study focuses only on reproductive toxicity.</p>
<p>Krug AK, Balmer N V, Matt F, Schönenberger F, Merhof D & Leist M (2013). Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. <i>Archives of Toxicology</i> 87: 2215-2231.</p>	<p>In this paper certain assay features were evaluated that can affect test specificity and some general procedures are suggested on how positive hits in complex biological assays may be defined. Organ-specific <i>in vitro</i> toxicity assays are often highly sensitive, but they lack specificity. Differentiating human LUHMES cells were used as a potential model for developmental neurotoxicity testing. Forty candidate toxicants were screened, and several hits were obtained and confirmed. Although the cells had a definitive neuronal phenotype, the use of a general cell death endpoint in these cultures did not allow specific identification of neurotoxicants. As an alternative approach, neurite growth was measured as an</p>

	<p>organ-specific functional endpoint. It was found that neurite extension of developing LUHMES was specifically inhibited by diverse compounds and these compounds reduced neurite growth at concentrations that did not compromise cell viability, and neurite growth was affected more potently than the integrity of developed neurites of mature neurons. A ratio of the EC50 values of neurite growth inhibition and cell death of >4 provided a robust classifier for compounds associated with a developmental neurotoxic hazard. The assay also identified compounds that accelerated neurite growth, such as the rho kinase pathway modifiers blebbistatin or thiazovivin. In summary, it was suggested that assays using functional endpoints (neurite growth) can specifically identify and characterize (developmental) neurotoxicants.</p>
<p>Robinson, J. F., Verhoef, A., Pennings, J. L. a, Pronk, T. E., & Piersma, A. H. (2012). A comparison of gene expression responses in rat whole embryo culture and in vivo: time-dependent retinoic acid-induced teratogenic response. <i>Toxicological Sciences</i> 126(1), 242–54.</p>	<p>In this paper, the effects of the model teratogen, all-trans retinoic acid (RA) in WEC and embryos <i>in vivo</i> were investigated. The transcriptomic approach was applied to monitor the effects of RA on gene expression in RA exposed and nonexposed rat embryos derived using WEC or <i>in vivo</i>. In addition, morphological changes of the embryos were studied. Across six time points the strong similarities in RA response at the gene and functional level were observed. The differences between models in the timing of RA-induced effects on genes related to embryonic development and RA metabolism were observed. The observations on the gene expression level were associated with specific differential morphological outcomes. This study supports the usefulness of the WEC to examine compound-induced molecular responses relative to <i>in vivo</i> embryos.</p>
<p>Schenk, B., Weimer, M., Bremer, S., Van der Burg, B., Cortvrindt, R., Freyberger, A., ... Schwarz, M. (2010). The ReProTect Feasibility Study, a novel comprehensive in vitro approach to detect reproductive toxicants. <i>Reproductive toxicology</i> 30(1), 200–18.</p>	<p>ReProTect is a project within the 6th European Framework Program which has developed alternative methods aimed to reduce or replace animal experimentation in the field of reproductive toxicology. In its final year, a ring trial, named the “Feasibility Study”, was conducted, in which 10 blinded chemicals with toxicologically well-documented profiles were analyzed by employing a test battery of 14 <i>in vitro</i> assays. EC₅₀ (half maximal effective concentration) or equivalent endpoints were determined and the test compounds were ranked relative to chemicals previously assayed in the tests of the battery. This comparative analysis together with a weight of evidence approach allowed a robust prediction of adverse effects on fertility and embryonic development of the 10 test chemicals <i>in vivo</i>. In summary, the vast majority of the predictions made based on the <i>in vitro</i> results turned out to be correct when compared to the whole</p>

	animal data. The procedure used here, a nearest neighbour analysis coupled with a weight of evidence approach, may guide future activities in the field of alternative toxicity testing.
--	--

11. Endocrine disruption relevant to human health and wildlife populations

Malgorzata Nepelska, Julien Burton, Marina Goumenou, Sharon Munn & Andrew Worth

11.1 What are endocrine disruptors and why are they of concern?

According to the World Health Organization (WHO), an “endocrine disrupter” is defined as “...an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations”, whereas a “potential endocrine disrupter” is “... an exogenous substance or mixture that possesses properties that might be expected to lead to endocrine disruption in an intact organism, or its progeny, or (sub)populations.” (WHO/IPCS, 2002)

Endocrine disruptors are of societal concern because of their adverse effects on humans, wildlife and the environment (EEA, 2012). For example, endocrine-related adverse human health effects include certain cancers, as well as effects on fertility and development (UNEP WHO, 2013).

The endocrine system can be considered to encompass all of the hormone-producing tissues that release hormones into the blood, which transports the hormones to target tissues where they elicit their physiological effects. The major functions of hormones are to regulate various functions including energy metabolism (i.e. energy storage, production, and utilisation); growth and development; and function of the immune and reproductive system.

In classical endocrinology, the endocrine system comprises multiple axes consisting of different organs and hormones with distinct regulatory functions, the three main axes being the hypothalamus-pituitary-gonad (HPG), the hypothalamic-pituitary-adrenal (HPA), and the hypothalamic-pituitary-thyroid (HPT) axes. However, as noted by Kortenkamp et al. (2012), scientific advances in our understanding of receptor signalling and molecular biology are continuously blurring the borders between the nervous system, immune system and endocrine system, which has the important implication that our understanding of the endocrine system and endocrine signalling “can therefore span from the classical definition of the endocrine system to one that encompasses any type of receptor-mediated signalling.”

This finely tuned system may be susceptible to perturbation by exogenous substances, Endocrine Active Substances (EAS), which may in some circumstances result in adverse effects, including effects on foetal development and disruption of homeostasis in the adult organism. The Endocrine Society proposes in a position statement from 2012 (Zoeller et al, 2012) that “the dose makes the poison”, since screening and testing for EDs and estimating potency require insight derived from principles of endocrinology according to which (endogenous) hormones:

1. coordinate the development and function of tissues in a highly integrated manner
2. act via receptors
 - a. receptors are tissue specific, most sensitive to the hormones at the low end of the dose response curve, and amplify the response;
 - b. affinity for receptors is distinct from overall *in vivo* potency;
 - c. response is dependent on hormone concentration, receptor affinity, receptor abundance and co-regulatory proteins;
 - d. binding to different receptors may occur as dose increases (receptor cross-talk), resulting in different responses from those seen at low doses;
3. act at very low concentrations, in the ppt-ppb range;
4. often result in non-monotonic dose responses due to multiple mechanisms and

5. produce effects that are life stage dependent, with a greater potential for permanent effects occurring following activity during development

Since hormones generally act at low doses and may activate their receptors in a non-linear fashion, there is concern that those EAS acting similarly to hormones may, initiate relevant effects depending on affinity to receptors also in a non-linear fashion, and some, depending on potency, may do so at very low concentrations. Most dose-response relationships follow the classical sigmoidal pattern, but the dose-relationship can also be more complex, including non-monotonic dose-response relationships (NMDR) (Vandenberg et al. 2012). NMDRs can be defined as a dose-response which changes direction from ascending to descending or vice versa and can occur at any part of the dose axis, producing, for example, U-shaped (with maximal responses observed at low and high doses) or inverted U-shaped (with maximal responses observed at intermediate doses) curves. Although, NMDRs are not specific to ED-related endpoints, if hormones generally act with NMDRs, it is considered not unlikely that EAS that mimic hormones may also exhibit NMDR behaviour. One of the concerns related to current assessment methodology is that the occurrence of NMDR at doses below those tested in standardised assays cannot be ruled out *a priori*, and thus there is uncertainty whether a safety threshold derived from a classical (sigmoidal) dose-response relationship is sufficiently protective (Zoeller et al., 2012). This uncertainty is compounded by the fact that the traditional animal tests might not include the most sensitive endocrine-relevant endpoints. According to EFSA's Scientific Committee (EFSA, 2013) the debate is evolving in the scientific community as to the existence and/or relevance of low-dose effects and NMDRs in (eco)toxicology in relation to endocrine disruption or other endpoints/modes of action, but still lacks consensus. More work needs to be conducted to agree the definitions of the respective terms, and in practical terms to consider whether or how it could impact upon risk assessment (i.e. assessment of dose response relationships for adverse effects) and testing strategies.

Another challenge related to the assessment of endocrine disruption is that the sensitivity of an organism to the action of hormones (and possibly EAS as well) varies throughout life, with increased sensitivity during critical periods of development. There are important differences in the maturity and functionality of the endocrine system between the pre- and post-natal life, including the lack of fully developed endocrine axes during foetal development. In addition, puberty, pregnancy and menopause are adult life stages which may be particularly sensitive to endocrine disruption. In this context a relevant issue is the possibility of delayed effects (delay between the exposure and the appearance of the adverse effect). For example, exposure to EAS in foetal life may lead to adverse effects during adulthood. Consequently, another concern related to current assessment methodology is that current tests are not capable of detecting the adverse effects of EAS that are triggered during sensitive windows of development and which may not be apparent until a later life stage.

11.2 Regulatory implications for the identification of endocrine disruptors

Already in the 90s, the Community Strategy for Endocrine Disruptors (EC, 1999) aimed at creating a list of substances requiring priority evaluation (“ED priority list”) for their role in endocrine disruption and to identify substances which could already be addressed under existing legislation, as well as addressing gaps in knowledge and specific cases of consumer use for special consideration. A candidate list of 553 chemicals was published in 2000, together with a series of actions proposed to further evaluate the role of these substances in endocrine disruption. This list of chemicals and associated toxicity data are being incorporated into a web-based database, the Endocrine Active Substances Information System (EASIS), developed by the JRC (Castello & Worth, 2011). The stated long-term goal of the Community Strategy was “to control substances having harmful effects on humans, wildlife and/or the environment”.

The need to identify EDs and to assess the risks they pose is reflected in various pieces of EU legislation, with a range of consequences in terms of risk management and the marketing of chemicals and their products (Table 11.1). To support the implementation of these pieces of EU legislation, efforts are ongoing to establish criteria and guidance for the identification and risk assessment of EDs. The key scientific considerations for identifying and characterising EDs have recently been proposed by the European Commission's Endocrine Disrupters Expert Advisory Group (ED EAG; Munn & Goumenou, 2013)¹ and by the EFSA Scientific Committee (EFSA, 2013). According to the ED EAG, for a substance to be identified as an ED, it must "demonstrate an adverse effect for which there is convincing evidence of a biologically plausible causal link to an endocrine disrupting mode of action of which the disruption is not a secondary consequence of other non-endocrine-mediated systemic toxicity." For the characterisation of EDs, the ED EAG agreed that other factors should be considered, including potency, severity, (ir)reversibility, and lead toxicity/critical effect.

11.3 Mechanistic understanding of endocrine disruption

The assessment of endocrine disruption poses a scientific and technological challenge since there are many different mechanisms of endocrine-mediated toxicity, and many EAS are associated with multiple mechanisms, due to the "promiscuity" of ligand-receptor binding, and "cross-talk" between hormone receptor pathways. Furthermore, EAS with similar mechanisms may work together in additive, synergistic or permissive mode to produce combined effects. These combination effects can occur at doses where each chemical individually is without detectable effects (Kortenkamp, 2007).

While our mechanistic understanding of EAS action is still incomplete, different mechanisms can in principle be distinguished according to their Molecular Initiating Events (MIEs), i.e. the first chemico-biological interaction that results in a perturbation of a biological pathway.

Receptor-mediated mechanisms are based on the ability of chemicals to act directly as ligands for hormone receptors, like oestrogen, androgen, and thyroid receptors. Nuclear hormone receptors are a class of proteins that act together with other proteins to regulate the expression of specific genes (Gronemeyer et al., 2004). These receptors are designed to bind small, lipophilic molecules (i.e. the native endogenous ligands such as steroid hormones) but such characteristics may be shared by other exogenous substances, including environmental contaminants.

Other mechanisms of disruption are based on interference with the synthesis, release, transport and metabolism of hormones. For example, the molecular targets of EAS include the cytochrome P450 enzymes in the steroid biosynthesis pathway (Sanderson, 2006). In addition, there is also growing evidence for the involvement of epigenetic mechanisms and their trans-generational consequences (Hung et al. 2010; Zhang & Ho 2010; Manikkam et al., 2013), although the implications of such mechanisms for human health are still poorly understood.

11.4 Status of methodology, including non-standard methods

A wide range of *in silico* models, *in vitro* and *in vivo* methods have been or are being developed to support the assessment of EDs, as reviewed in various reports (Lo Piparo &

¹ The Endocrine Disrupters Expert Advisory Group (ED EAG) was established in November 2011 as a sub-group of the ad hoc group of Commission Services, EU Agencies and Member States for the Community Strategy on Endocrine Disrupters. The main role of the ED EAG has been to provide detailed reflections on scientific issues relevant to endocrine disrupting substances, not specific to any regulatory framework, including advice/orientation on scientific criteria for the identification of endocrine disrupting substances.

Worth, 2010; EFSA, 2013) and by the OECD (OECD 2012a; OECD 2012b). However, no single model or assay is likely to provide all the information needed to label a substance as an ED, since information on both the mode of action and the adverse effect will be required, if recommendations from the EFSA and the ED EAG are followed (EFSA, 2013; Munn & Goumenou, 2013).

The OECD Conceptual Framework (CF) for the testing and assessment of endocrine disruptors provides a structured guide to the standardised test methods available or proposed as OECD Test Guidelines (OECD, 2012a). According to OECD CF (Table 11.2), information on endocrine activity can be obtained from existing information, read-across, and *in silico* tools (collectively referred to as Level 1), as well as *in vitro* (Level 2) and *in vivo* (Level 3) screening assays providing mechanistic information. Since a prerequisite for an EAS to be regarded as an ED is likely to be the identification of an endocrine-mediated adverse effect, *in vivo* test methods based on apical endpoints to demonstrate the adverse pathological or functional effect in the intact organism are also included (Levels 4 and 5). It should be noted that the OECD CF is not intended to be a testing strategy, but rather a list of potentially useful methods and data sources. Most of the level 2, 3 and 4 endocrine specific methods in the OECD CF are not explicitly required for the assessment of chemicals under REACH, PPPR or the BPR. For assessment of effects on wildlife populations this also applies to the methods indicated under level 5.

Level 1 of the OECD CF consists of all available information, including physicochemical properties, data from standardised and non-standardised tests, epidemiological and field studies, and *in silico* predictions.

Given the wide variety of mechanisms by which EDs can act, it is unrealistic to capture all endocrine disrupting activity in a single *in silico* model. These models have therefore focused on the prediction of endocrine-related adverse outcomes (such as reproductive and developmental toxicity) or on the binding and subsequent activation or inactivation of receptors associated with these adverse outcomes (Lo Piparo & Worth, 2010; Cronin & Worth, 2008). The most popular target for modelling has been the Oestrogen Receptor (ER), with methods including both ligand-based and receptor-based modelling (Tsakovska et al., 2011). Taking advantage of the numerous available 3D structures of the ER receptor (several high resolution structures of the α , β , and dimer forms co-crystallized with various ligands), docking simulations and 3D QSAR methods like CoMFA (Comparative Molecular Field Analysis) have been used. These methods simulate key MIEs associated with endocrine disruption. An alternative modelling approach has been to apply machine learning algorithms (e.g. linear regression, decision trees, nearest neighbours, neural networks, support vector machines) to molecular descriptors (Novic & Vracko, 2010). These statistical methods, based exclusively on the ligand properties, offer potentially useful classifiers, but are often lacking mechanistic interpretability. To a lesser extent, the androgen receptor and thyroid receptor have also been modelled. A promising and more recent development has been to combine automated docking simulations with QSAR (Vedani et al., 2012). The advantages and limitations of the main types of *in silico* methods are given in Table 11.3.

Level 2 of the CF currently consists of *in vitro* methods for detecting estrogenic, androgenic and steroidogenic activity, whereas *in vitro* test guidelines based on thyroid activity are currently lacking. Examples of recently developed *in vitro* test guidelines include:

- a) OECD TG 455, a performance based test guideline for estrogenic (agonist) activity via ER transactivation;
- b) OECD TG 456, a human adenocarcinoma cell line (H295R) for effects on steroidogenesis; and

- c) OECD TG 457, an ER transactivation assay based on a human ovarian adenocarcinoma cell line (BG1Luc4E2) to identify both estrogenic and anti-estrogenic activity. Other methods are undergoing validation by EURL ECVAM, including the MELN assay for ER agonism/antagonism (study in progress) and several assays for AR agonism/antagonism (study in planning phase). The advantages and limitations of the main types of *in vitro* test are given in Table 11.4.

Level 3 consists of *in vivo* screening assays that provide insight into chemical interactions with single selected endocrine mechanism(s) / signalling pathways.

Levels 4 and 5 consist of assays that assess *in vivo* adverse apical outcomes including endocrine-relevant endpoints. These assays provide insight into chemical interactions with multiple endocrine signalling pathways or endpoints. However, effects can be sensitive to more than one mechanism and may be due to non-endocrine mechanisms.

Level 5 consists of whole-organism assays that are designed to provide more comprehensive data on adverse effects on endocrine-relevant endpoints over more extensive parts of the life cycle of the organism than level 4.

There are many additional activities ongoing contributing to the development of methodologies for the assessment of endocrine disrupters, a number of which are being conducted under the auspices of OECD. Notably, the OECD Detailed Review Paper on “the State of the Science on novel *in vitro* and *in vivo* screening and testing methods and endpoints for evaluating endocrine disruptors” (OECD, 2012b) provides an evaluation of further endocrine pathways and targets for disruption, along with assays that show promise in evaluating the target in a screening and testing programme, in order to consider such assays for validation to complement the existing OECD Test Guidelines. In particular, an OECD Expert Group is currently focusing on the elucidation of pathways linked to disruption of thyroid hormone and the hypothalamic-pituitary-thyroid (HPT) axis, identifying the most promising assays capable of probing various modes of action to which could be brought forward for validation.

One of the critical missing components in the translation of *in vitro* data to the *in vivo* situation is knowledge of absorption, distribution, metabolism and excretion (ADME) in the intact organism. Consequently knowledge of the metabolic fate of the parent compound and delivery of the active moiety to the relevant tissue has also to be developed before it will be possible to accurately predict endocrine disruption *in vivo* (Jacobs et al., 2013). In relation to metabolism, the characterising of cells used in such assays for metabolic capacity and consideration of introducing additional metabolising systems into current *in vitro* assays is also being taken up within the OECD.

11.5 Adverse Outcome Pathways (AOPs)

Adverse Outcome Pathways is a concept in which existing knowledge concerning the linkage between a direct molecular initiating event and an adverse outcome is portrayed at a level of biological organisation relevant to risk assessment. Once an adverse outcome pathway (AOP) has been elucidated for one chemical whereby the series of events from MIE through a series of key events to an adverse outcome has been identified for one chemical or group of chemicals it may become possible to associate other chemicals to the group not only on the basis of structure but also in relation to shared modes of action.

Although not specific to endocrine disrupters, it can be envisaged that once the linkages between upstream key events which can be measured *in vitro* can be quantitatively linked to the likelihood of an adverse effect occurring *in vivo* such evidence could be used in hazard

identification of EDs as well as providing points of departure for risk assessment. The development of AOPs for a number of adverse outcomes via endocrine modes of action is currently in progress at OECD. Examples include the AOPs linking Aromatase Inhibition, Androgen Receptor Agonism, Oestrogen Receptor Antagonism, and Steroidogenesis Inhibition, to Impaired Reproduction in Small Repeat-Spawning Fish Species. Further information is available from the OECD (<http://www.oecd.org/env/ehs/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm>).

One of the critical missing components in the translation of *in vitro* data to the *in vivo* situation is knowledge of absorption, distribution, metabolism and excretion (ADME) in the intact organism. Consequently knowledge of activation or deactivation of the parent compound and delivery of the active moiety to the relevant tissue has also to be developed before it will be possible to accurately predict endocrine disruption *in vivo* (OECD 2008; Jacobs et al, 2013). In relation to metabolism, the characterising of cells used in *in vitro* assays for metabolic capacity and consideration of introducing exogenous metabolising systems into current *in vitro* assays is also being actively pursued at OECD level. In addition an OECD Expert Group is currently focusing on the elucidation of pathways linked to disruption of thyroid hormone and the hypothalamic-pituitary-thyroid (HPT) axis, identifying the most promising assays capable of probing various modes of action which could be brought forward for validation. The OECD Detailed Review Paper No.178 (OECD, 2012b) describes a number of other endocrine-related pathways and available assays to measure modulation of such pathways which may be suitable for further development and validation, and some member countries are already proposing projects to develop assays for other endocrine pathways beyond estrogen, androgen, and thyroid-related pathways.

11.6 Conclusions

The policy need to identify EDs within REACH and the Biocidal Products Regulation (as well as a number of other chemical control-related pieces of legislation including the Plant Protection Products Regulation) drives the need to develop assays that can identify not only the toxicity of a chemical but also the mode of action leading to toxicity. In order to proceed with risk assessment quantitative information on dose-response relationships will also be required.

According to the opinion of the ED EAG, also supported by EFSA's Scientific Committee, the elements for identification of an endocrine disrupter are demonstration of an adverse effect for which there is convincing evidence of a biologically plausible link to an endocrine disrupting mode of action and for which disruption is not a secondary consequence of other non-endocrine mediated systemic toxicity. The assays and tools specified in levels 1 to 5 of the OECD CF for the testing and assessment of endocrine disrupters provides a suite of assays that allows the identification of an adverse effect in the apical *in vivo* studies and the indication of some specific endocrine modes of action via disruption of oestrogen or androgen signalling, inhibition of steroidogenesis and some *in vivo* biomarkers of endocrine disruption. At the moment these mechanistic assays are foreseen to be used either in a screening and prioritisation context to identify potential endocrine disrupters, which can be followed with appropriate *in vivo* testing to identify plausibly linked adverse effects, or to investigate adverse effects already identified in *in vivo* studies that are suspected of being caused by an endocrine disrupting mode of action. Further work is needed to establish integrated assessment strategies not reliant on animal testing to meet the regulatory definition of an ED. Therefore, it is currently unlikely that *in vivo* testing could be avoided in most cases apart from where strong arguments for read-across could be justified.

A wide range of activities are in progress which can contribute towards the goal of developing alternative non-animal approaches for use in the identification of endocrine disrupters, but

considerable work is still required to optimise individual methods and combine them in a manner that could adequately model the complexity of endocrine disruption. Since knowledge of the underlying mechanisms of ED action is still incomplete, and the battery of suitable *in silico* models and *in vitro* tests inadequate, it is not possible to determine, with a high degree of certainty, whether a substance is an ED without performing appropriate *in vivo* studies. Furthermore, even currently available *in vivo* assays might result in an underestimation of the full extent of the risks associated with EDs, due to factors such as the sensitivity of the animal model in relation to humans, the possible lack of inclusion of the most sensitive endpoints during the most sensitive life stages in current *in vivo* assays and particularly the possibility of missing delayed and potentially serious effects.

The OECD CF provides a useful way of categorising different (animal and non-animal) models according to the type of information they provide. Furthermore, the AOP approach provides a useful means of organising knowledge on the mechanistic pathways of endocrine disruption, thereby forming a rational basis for developing integrated assessment approaches and building weight-of-evidence arguments. For example, computational and *in vitro* models that provide information on ligand-receptor interactions could be used to establish the likelihood of ED-related MIEs. This information could then be combined with predictions of downstream events, including the final adverse outcome (AO). If the AOP is well established, evidence for a “plausible causal link” between the MIE and the AO will already be available; conversely, if the AOP is tentative, further information (either from the literature or generated by *ad hoc* studies) would be needed to support the conclusion that a substance is an ED.

11.7 References

Castello P & Worth A (2011). Information sources and databases on Endocrine Active Substances. JRC Technical Note CT 31854. 28 pp. Available from: http://ihcp.jrc.ec.europa.eu/our_activities/food-cons-prod/endocrine_disrupters/eas_database/info-sources-databases-endocrine-active-substances

Cronin M & Worth A (2008). (Q)SAR for predicting effects relating to reproductive toxicity. QSAR & Combinatorial Science 27: 91–100.

EC (1999). Community Strategy for Endocrine Disrupters - a range of substances suspected of interfering with the hormone systems of humans and wildlife. European Commission. Available from: http://europa.eu/legislation_summaries/internal_market/single_market_for_goods/chemical_products/21277_en.htm

EEA (2012). The impacts of endocrine disruptors on wildlife, people and their environment – the Weybridge +15 report. European Environment Agency. Available from: <http://www.eea.europa.eu/publications/the-impacts-of-endocrine-disruptors>

EFSA (2013). Scientific Opinion on the hazard assessment of endocrine disruptors: scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA Journal 2013;11(3):3132. European Food Safety Authority. Available from: <http://www.efsa.europa.eu/en/efsajournal/doc/3132.pdf>

Gronemeyer H, Gustafsson JA & Laudet V (2004). Principles for Modulation of the Nuclear Receptor Superfamily. Nature Reviews. Drug Discovery 3: 950–64.

Hung CH, Yang SN, Kuo PL, Chu YT, Chang HW, Wei WJ, Huang SK & Jong YJ. (2010). Modulation of Cytokine Expression in Human Myeloid Dendritic Cells by Environmental Endocrine-disrupting Chemicals Involves Epigenetic Regulation. Environmental Health Perspectives 118: 67–72.

Jacobs MN, Laws S, Willett K, Schmeider P, Odum J & Bovee TF (2013) In vitro metabolism and bioavailability test for endocrine active substances. What is needed for regulatory purposes? *Altex* 30: 331-351.

Kortenkamp A (2007). Ten Years of Mixing Cocktails: a Review of Combination Effects of Endocrine-disrupting Chemicals. *Environmental Health Perspectives* 115 Suppl (December) 98-105.

Kortenkamp A, Backhaus T & Faust M (2009). State of the Art Report on Mixture Toxicity. Final Report of EU Project Contract 070307/2007/485103/ETU/D.1. 391 pp. Available from: http://ec.europa.eu/environment/chemicals/pdf/report_Mixture%20toxicity.pdf

Kortenkamp A, Martin, O, Faust M, Evans R, McKinlay R, Orton F & Rosivatz E (2012). State of the art assessment of endocrine disrupters. DG Environment project contract number 070307/2009/550687/SER/D3.

Lo Piparo E & Worth A (2010). Review of QSAR Models and Software Tools for Predicting Developmental and Reproductive Toxicity. JRC Scientific and Technical Report EUR 24522 EN. Available from: <http://publications.jrc.ec.europa.eu/repository>

Novic M & Vracko M (2010). QSAR Models for reproductive Toxicity and Endocrine Disruption activity. *Molecules* 15: 1987-1999.

OECD (2012a). Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption. Series on Testing and Assessment No 150. ENV/JM/MONO(2012)22. Organisation for Economic Cooperation and Development, Paris, 524 pp. Available from: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282012%2922&d_oqlanguage=en

OECD (2012b). Detailed Review Paper on the State of the Science on Novel In vitro and In vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. Series on Testing and Assessment No 178. ENV/JM/MONO(2012)23. Organisation for Economic Co-operation and Development, 213 pp. Available from: <http://www.oecd.org/env/ehs/testing/49002244.pdf>

Mohan M, Tracey R, Guerrero-Bosagna C & Skinner MK (2013). Plastics Derived Endocrine Disruptors (BPA, DEHP and DBP) Induce Epigenetic Transgenerational Inheritance of Obesity, Reproductive Disease and Sperm Epimutations. *PloS One* 8 (1) (January).

Munn S & Goumenou M (2013). Key scientific issues relevant to the identification and characterisation of endocrine disrupting substances. Report of the Endocrine Disruptors Expert Advisory Group. JRC Scientific and Policy Report EUR 25919 EN. Publications Office of the European Union, Luxembourg. Available from: <http://publications.jrc.ec.europa.eu/repository>

Sanderson JT (2006). The Steroid Hormone Biosynthesis Pathway as a Target for Endocrine-disrupting Chemicals. *Toxicological Sciences* 94: 3–21.

Tsakovska I, Pajeva I, Alov P & Worth A (2011). Recent advances in the Molecular Modelling of Estrogen Receptor-Mediated Toxicity. *Advances in Protein Chemistry and Structural Biology* 85: 217-251.

UNEP WHO (2013). State of the science of endocrine disrupting chemicals – 2012. Available from: http://unep.org/pdf/9789241505031_eng.pdf

Vedani A, Dobler M & Smiesko M (2012). VirtualToxLab - A platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicology & Applied Pharmacology* 261: 142–153.

Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs, DR Lee DH, Shioda T, et al. (2012). Hormones and Endocrine-disrupting Chemicals: Low-dose Effects and Nonmonotonic Dose Responses. *Endocrine Reviews* 33: 378–455.

WHO/IPCS (2002). Global Assessment of the State-of-the-science of Endocrine Disruptors. World Health Organization/International Programme on Chemical Safety. WHO/PCS/EDC/02.2, 180 pp. Available from: http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en

Zhang X & Ho SM (2010). Epigenetics Meets Endocrinology. *Journal of Molecular Endocrinology* 46: R11–R32.

Zoeller, R. T., Brown, T. R., Doan, L. L., Gore, A. C., Skakkebaek, N. E., Soto, A. M., ... Vom Saal, F. S. (2012). Endocrine-disrupting chemicals and public health protection: a statement of principles from The Endocrine Society. *Endocrinology* 153(9): 4097–4110.

Table 11.1. The treatment of EDs in REACH, PPPR, BPR and the cosmetic products regulation

Legislation	ED-related provisions
REACH Regulation (EC No 1907/2006), Art. 57(f)	Stipulates that substances subject to authorisation may include ...those having endocrine disrupting properties ...for which there is scientific evidence of probable serious effects to human health or the environment which give rise to an equivalent level of concern as Carcinogens, Mutagens and Reproductive and Developmental Toxicants (CMR) or Persistent, Bioaccumulative and Toxic (PBTs) or very Persistent and very Bioaccumulative (vPvBs) substances and which are identified on a case by case basis in accordance with Article 59 (a review of the authorisation procedure with regard to endocrine disruptors is required by 1 June 2013 (cf. Article 138(7) ²).
Regulation (528/2012) on Biocidal Product (BPR)	<p>indicates that active substances shall not be approved if they are considered as having endocrine-disrupting properties that may cause adverse effects in humans or which are identified in accordance with Article 57(f) and 59(1) of REACH as having endocrine disrupting properties. No later than 13 December 2013 the Commission shall adopt scientific criteria for the determination of endocrine disrupting properties³</p> <p>Pending the adoption of these criteria, substances that are or have to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as</p> <ul style="list-style-type: none"> • carcinogenic category 2 and toxic for reproduction category 2, shall be considered to have endocrine disrupting properties. • toxic for reproduction category 2 and which have toxic effects on the endocrine organs, may be considered to have such endocrine disrupting properties.

² As of 4 April 2014 review not completed

³ As of 4 April 2014 criteria not yet adopted

Legislation	ED-related provisions
Regulation on Plant Protection Products (EC No 1107/2009) (PPPR)	<p>stipulates that an active substance, safener or synergist shall only be approved if, ...it is not considered to have endocrine disrupting properties that may cause adverse effects in humans or non-target organisms, unless the exposure is ...negligible. Furthermore, the Commission is mandated to present a draft of the measures concerning specific scientific criteria for the determination of endocrine disrupting properties by 13 December 2013⁴.</p> <p>Pending the adoption of these criteria, substances that are or have to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as</p> <ul style="list-style-type: none"> • carcinogenic category 2 and toxic for reproduction category 2, shall be considered to have endocrine disrupting properties. • toxic for reproduction category 2 and which have toxic effects on the endocrine organs, may be considered to have such endocrine disrupting properties.
Cosmetic Products Regulation (EC No 1223/2009), Article 15 (4).	states that as soon as Community or internationally agreed criteria for identifying substances with endocrine disrupting properties are available, or at the latest on 11 January 2015, the Commission shall review this Regulation with regard to substances with endocrine-disrupting properties.

⁴ As of 4 April 2014 criteria not yet proposed by COM

Table 11.2. OECD Revised Conceptual Framework for Testing and Assessment of Endocrine Disrupters

OECD Conceptual Framework	Mammalian and non-mammalian toxicology	
<p>Level 1. Existing information and non-test information</p>	<ul style="list-style-type: none"> • Physical & chemical properties, e.g., MW reactivity, volatility, biodegradability • All available (eco)toxicological data from standardized or non-standardized tests • Read across, chemical categories, QSARs and other in silico predictions, and ADME model predictions 	
<p>Level 2. <i>In vitro</i> assays providing data about selected endocrine mechanism(s) / pathway(s) (Mammalian and non mammalian methods)</p>	<ul style="list-style-type: none"> • Estrogen or androgen receptor binding affinity • Estrogen receptor transactivation (OECD TG 455) • Androgen or thyroid transactivation (If/when TGs are available) • Steroidogenesis in vitro (OECD TG 456) • MCF-7 cell proliferation assays (ER antagonist) • Other assays as appropriate 	
	Mammalian Toxicology	Non-mammalian Toxicology
<p>Level 3. <i>In vivo</i> assays providing data about selected endocrine mechanism(s) / pathway(s)¹</p>	<ul style="list-style-type: none"> • Uterotrophic assay (OECD TG 440) • Hershberger assay (OECD TG 441) 	<ul style="list-style-type: none"> • Xenopus embryo thyroid signalling assay (When/if TG is available) • Amphibian metamorphosis assay (OECD TG 231) • Fish reproductive screening assay (OECD TG 229) • Fish screening assay (OECD TG 230) • Androgenized female stickleback screen (GD 140)
<p>Level 4. <i>In vivo</i> assays providing data on adverse effects on endocrine relevant endpoints²</p>	<ul style="list-style-type: none"> • Repeated dose 28-day study (OECD TG 407) • Repeated dose 90-day study (OECD TG 408) • 1-generation reproduction toxicity study (OECD TG 415) • Male pubertal assay (GD 150)³ • Female pubertal assay (GD 150)³ • Intact adult male endocrine screening assay (GD 150) • Prenatal developmental toxicity study (OECD TG 414) • Chronic toxicity and carcinogenicity studies (OECD TG 451-453) • Reproductive screening test (OECD TG 421 if enhanced) • Combined 28-day/reproductive screening assay (OECD TG 422 if enhanced) 	<ul style="list-style-type: none"> • Fish sexual development test (Draft OECD TG 234) • Fish reproduction Partial Lifecycle Test (when/If TG is Available) • Larval amphibian growth & development assay (when TG is available) • Avian reproduction assay (OECD TG 206) • Mollusc partial lifecycle assays (when TG is available)⁴ • Chironomid toxicity test (TG 218-219)⁴ • Daphnia reproduction test (with male induction) (OECD TG 211)⁴ • Earthworm reproduction test (OECD TG 222)⁴ • Enchytraeid reproduction test (OECD TG

OECD Conceptual Framework	Mammalian and non-mammalian toxicology	
	<ul style="list-style-type: none"> • Developmental neurotoxicity (OECD TG 426) 	220) ⁴ <ul style="list-style-type: none"> • Sediment water lumbriculus toxicity test using spiked sediment (OECD TG 225)⁴ • Predatory mite reproduction test in soil (OECD TG 226)⁴ • Collembolan reproduction test in soil (OECD TG 232)⁴
Level 5. <i>In vivo</i> assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism ²	<ul style="list-style-type: none"> • Extended one-generation reproductive toxicity study (OECD TG 443)⁵ • 2-Generation reproduction toxicity study (OECD TG 416 most recent update) 	<ul style="list-style-type: none"> • Fish lifecycle toxicity test (FLCTT) (when TG is available) • Medaka multigeneration test (MMGT) (when TG is available) • Avian 2 generation reproductive toxicity assay (when TG is available) • Mysid lifecycle toxicity test (when TG is available)⁴ • Copepod reproduction and development test (when TG is available)⁴ • Sediment water chironomid life cycle toxicity test (OECD TG 233)⁴ • Mollusc full lifecycle assays (when TG is available)⁴ • Daphnia multigeneration assay (if TG is available)⁴

Footnotes

- 1) Some assays may also provide some evidence of adverse effects.
- 2) Effects can be sensitive to more than one mechanism and may be due to non-ED mechanisms.
- 3) Depending on the guideline/protocol used, the fact that a substance may interact with a hormone system in these assays does not necessarily mean that when the substance is used it will cause adverse effects in humans or ecological systems.
- 4) At present, the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non- EDs. Those in Level 4 are partial lifecycle tests, while those in Level 5 are full- or multiple lifecycle tests.
- 5) The new EOGRT study (OECD TG 443) is preferable for detecting endocrine disruption because it provides an evaluation of a number of endocrine endpoints in the juvenile and adult F1, which are not included in the 2-generation study (OECD TG 416) adopted in 2001

Notes to the OECD Revised Conceptual Framework

- Note 1: Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information and needs for testing and assessment.
- Note 2: The assessment of each chemical should be made on a case by case basis, taking into account all available information.

Note 3: The framework should not be considered as all inclusive at the present time. At levels 2, 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included.

Table 11.3. Advantages and limitations of *in silico* methods for endocrine disruptors

Type of Method	Advantages	Limitations
QSAR and machine learning	<ul style="list-style-type: none"> • High accuracy statistical modelling. • Defined applicability domain. 	<ul style="list-style-type: none"> • Some machine learning methods and molecular descriptors lack mechanistic interpretation. • Choice of multiple parameters (algorithm, descriptors, ...)
Docking	<ul style="list-style-type: none"> • Simulation of the ligand/protein complex. • In depth mechanistic interpretation. 	<ul style="list-style-type: none"> • Need for receptor crystallographic structure. • Non-automated procedure may be time consuming.
Structural alerts	<ul style="list-style-type: none"> • Simple and quick evaluation of toxicity 	<ul style="list-style-type: none"> • Generalisation may be debatable (lots of exceptions)

Table 11.4. Advantages and limitations of *in vitro* tests for endocrine disruptors

Type of Method	Advantages	Limitations
Transactivation reporter assay	<ul style="list-style-type: none"> • provide quantitative & functional information • commercially available for many receptors • reporter assays could serve as a screening assay to discern a potential anchoring molecular event that would trigger assessment along the relevant adverse outcome pathway • known performance capabilities, OECD TG 455 ER transactivation assay describes several assays • overall recommended for interaction with nuclear receptors 	<ul style="list-style-type: none"> • some disrupting chemicals may interact with different nuclear receptors and would be missed if only one receptor or DNA array is employed, hence initial screening assays need to cover multiple targets within the axes of the endocrine system.
Receptor-binding assays	<ul style="list-style-type: none"> • Species-specific membrane binding experiments in conjunction with e.g. ER binding or transactivation reporter assays (e.g., OECD TG 455) would provide a complete molecular assessment of chemical-receptor interactions that may serve as the initiating event in the e.g. estrogen adverse outcome pathway • improved sensitivity achieved by including also co-activators 	Receptor binding alone does not provide information on activation or inactivation of the receptor
DNA/protein binding assays	<ul style="list-style-type: none"> • capacity to detect chemicals that are "not direct receptor ligands" 	
Cell-based microarrays	<ul style="list-style-type: none"> • capture changes in gene expression for massive numbers of genes and enable the simultaneous analyses of components along the signalling pathway, and its 	<ul style="list-style-type: none"> • cost and time intensive (assays require the construction of the arrays, many are commercially available, but are relatively

Type of Method	Advantages	Limitations
	<p>product</p> <ul style="list-style-type: none"> • fingerprint of the analysed compound • overall potential ability to assess chemical impact on multiple endocrine signalling pathways simultaneously however the approach may not be sufficiently developed for routine, validated use at this time 	<p>expensive)</p> <ul style="list-style-type: none"> • analyses of the mass of data is complex • often lack of reproducibility • significant variability in gene responses attributed to cell type, agonist, arbitrary selection of threshold response levels, and lack of intra-experiment replication
<p>General cell based assays</p>	<ul style="list-style-type: none"> • cell based assays are relatively simple to perform, time and cost effective, and cell propagation methods standardized and able to accommodate high throughput and can identify endocrine toxicants that interact directly with hormone receptors • <i>in vitro</i> systems offer good reproducibility because effects are measured using the same cellular background, however, inter-laboratory collaborations are necessary for standardization and validation • the lower biological complexity of <i>in vitro</i> systems compared to the situation <i>in vivo</i> makes the data interpretation more straightforward • <i>in vitro</i> methods provide the information about molecular initiating events 	<ul style="list-style-type: none"> • additional complexities within relevant cell-types by assaying the normal function of the cells as related to the endocrine signalling pathway under investigation • some assays require the isolation of primary cells from animals, and thus are not strictly non-animal methods • species extrapolation problems (mouse models may not be always relevant for humans) • many established cell lines available (including human-derived) but may not behave as normal primary cells and thus relevance to <i>in vivo</i> situation still unclear • <i>in vitro</i> studies in cell systems are predominantly limited to assessment of parent compound examined as e.g. receptor transactivation only reflects primary ligand binding unless cells are “metabolically” active • the interactions demonstrated between HPA, HPG, and HPT axes, and the immune system, indicate that generally <i>in vitro</i> assays cannot be considered definitive for screening of chemicals for potential endocrine disruption since they do not capture these interactions

Chapter 11. Table of References with Notes

Key reports/guidance by governmental, inter-governmental, industrial and academic organisations	
<p>Castello P & Worth A (2011). Information sources and databases on Endocrine Active Substances. JRC Technical Note CT 31854. 28 pp. Available from: http://ihcp.jrc.ec.europa.eu/our_activities/food-cons-prod/endocrine_disrupters/eas_database/info-sources-databases-endocrine-active-substances</p>	<p>This review contains the status of (IT) information sources worldwide related with Endocrine Active Substances (EAS) up to January 2011. This work was based on a combination of literature and internet research, as well as on direct contacts with different organisations which host or use information systems directly or potentially related with EAS. Three types of information resources were analysed: existing databases with EAS data, software tools and computer models, and systems and/or conceptual approaches not specifically focusing on EAS, which may nevertheless inspire the design of an EAS Web Portal. The review aimed at bringing forth the conceptual approaches looking for useful functionalities for an EAS database and Web Portal.</p>
<p>EC (1999). Community Strategy for Endocrine Disrupters - a range of substances suspected of interfering with the hormone systems of humans and wildlife. European Commission. Available from: http://europa.eu/legislation_summaries/internal_market/single_market_for_goods/chemical_products/121277_en.htm</p>	<p>In this communication, the Commission proposes a strategy comprising short, medium and long-term action to deal with the problem of endocrine disrupters which threatens health and the environment. These actions are based on an assessment of the likely timeframe within which results can be achieved, i.e. 1-2 years for short-term actions, 2-4 years for medium-term actions and more than 4 years for long-term action.</p>
<p>ECETOC (2009). Guidance on Identifying Endocrine Disrupting Effects. Technical Report No 106, 133 pp.</p>	<p>This report reviews and summarises existing definitions for an endocrine disrupter as well as the test methods currently available to identify endocrine activity and/or toxicity. Specific scientific criteria for the determination of endocrine disrupting properties that integrate information from both regulatory (eco)toxicity studies and mechanistic/screening studies are proposed. These scientific criteria rely on the nature of the adverse effects detected in regulatory (eco)toxicity study(ies) that give concern for endocrine toxicity and the description/understanding of the mode of action of toxicity which scientifically support and explain the adverse effects. The criteria are presented in the form of flow charts for assessing relevant effects for both human and environmental species, using example substances. In addition since all chemicals having endocrine disrupting properties may not represent the same hazard, an element or assessment of potency is also proposed to discriminate chemicals of high concern from those of lower concern.</p>
<p>ECETOC (2011). Risk Assessment of Endocrine Disrupting Chemicals. Workshop Report No 21, 32 pp.</p>	<p>This report documents the outcome of a workshop organised by ECETOC to discuss the "Risk Assessment of Endocrine Disrupting Chemicals" held in Florence in May 2011. The experts from academia, regulatory bodies and industry discussed approaches for the risk assessment of endocrine disrupting chemicals. The workshop consisted of a series of invited presentations covering area of human and environmental safety. National initiatives and developments to define and test criteria for the identification of endocrine disrupting chemicals were presented. The presentations were followed by discussion sessions, which addressed four specific themes</p>

<p>EFSA (2013). Scientific Opinion on the hazard assessment of endocrine disruptors: scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA Journal 2013;11(3):3132. European Food Safety Authority. Available from: http://www.efsa.europa.eu/en/efsajournal/doc/3132.pdf</p>	<p>considered from both toxicological and ecotoxicological perspectives.</p> <p>This document contains key considerations on the scientific criteria to distinguish between; EDs and other groups of substances with different modes of action, physiological modulation and adverse effects on humans and on the ecosystem as a result of exposure to endocrine active substances (EASs) Additionally, provides a review existing test methods and discuss their appropriateness for the identification and characterisation of effects mediated by EASs. EFSA's opinion also briefly discusses several issues related to the testing of chemicals which are not unique to endocrine disruption: these include 'windows of susceptibility' (critical periods of development such as conception, pregnancy, infancy, childhood and puberty, when the body may be more sensitive to chemicals possibly increasing the likelihood of harmful effects in the short-term or later in life); and other matters related to chemical toxicology which are the subject of on-going scientific debate, such as low dose effects, non-monotonic (e.g. U-shaped) dose-response curves and exposures to multiple chemical substances. Moreover the authority recommend as a follow-up activity further work to clarify in a broader context how these various aspects could impact on current risk assessment approaches and testing strategies for any chemical substance.</p>
<p>OECD (2012a). Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Environmental Health and Safety Publications, Series on Testing and Assessment n°150, Organisation for Economic Cooperation and Development, Paris. http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282012%2922&doclanguage=en</p>	<p>The objectives of this Guidance Document are to:</p> <ul style="list-style-type: none"> • support regulatory authorities' decisions on the hazard of specific chemicals and toxicologically-relevant metabolites when they receive test results from a Test Guideline (TG) or draft TG for the screening/testing of chemicals for endocrine disrupting properties. The contexts for these decisions will vary, depending on local legislation and practice, so the advice is worded in such a way as to permit flexible interpretation. • provide guidance on how to interpret the outcome of individual tests and how to increase evidence on whether or not a substance may be an ED. Testing strategies or guidance on interpretation from a suite of tests are not given <p>Hazard assessment methods in this document are arranged in a two step process, with the intention of minimising animal testing globally through application of the 3Rs (Replace, Reduce and Refine the use of laboratory animals in testing):</p> <ul style="list-style-type: none"> • Use of a harmonised framework for assessing test results together with existing information on likely or known hazards should avoid unnecessary animal testing. Recommendation of a test method that may be performed if regulatory authorities need more evidence. The test method is defined precisely to facilitate the Mutual Acceptance of Data and to

	avoid unnecessary duplication of testing. The recommended test method will utilise non-animal tests where possible although a few alternative scenarios are considered depending upon existing information.
US EPA (2012a). Endocrine Disruptor Screening Program: Universe of Chemicals and General Validation Principles. November 2012. 17 pp. Available from: http://www.epa.gov/endo/pubs/edsp_chemical_universe_and_general_validations_white_paper_11_12.pdf	This document was developed by the EPA to provide guidance to staff and managers regarding the universe of chemicals and general validation principles for consideration of computational toxicology tools for chemical prioritization. It explores the chemical space covered by endocrine disrupting substances and discusses the appropriate way to apply computational tools.
US EPA (2012b). Endocrine Disruptor Screening Program: Universe of Chemicals. November 2012. 177 pp. Available from: http://www.epa.gov/endo/pubs/edsp_chemical_universe_list_1_12.pdf	Supporting materials of US EPA (2012a). Contains a list of +/-10000 chemicals suspected to show endocrine disrupting properties
US EPA (2013). Prioritizing the Universe of Endocrine Disruptor Screening Program (EDSP) Chemicals Using Computational Toxicology Tools. Meeting minutes and materials. Available from: http://www.epa.gov/scipoly/sap/meetings/2013/012913meeting.html	Meeting minutes and additional materials (panel members, supporting materials, ...) of the meeting "Prioritizing the Universe of Endocrine Disruptor Screening Program (EDSP) Chemicals Using Computational Toxicology Tools held" on January 29-31 2013.
Munn S and Goumenou M (2013). Key scientific issues relevant to the identification and characterisation of endocrine disrupting substances. Report of the Endocrine Disrupters Expert Advisory Group. JRC Scientific and Policy Report EUR 25919 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository	This report represents the opinions of The Endocrine Disrupters Expert Advisory Group (ED EAG) which is a sub-group of the ad hoc group of Commission Services, EU Agencies and Member States for the Community Strategy on Endocrine Disrupters. The scope of this report is to capture the experts' opinions on key scientific issues relevant to the identification of endocrine disrupting substances (EDs) in order to support the ad-hoc group discussion and the Commission's decisions on the establishment of horizontal criteria for the identification of EDs for use in different regulatory contexts.
Munn S and Goumenou M (2013). Thresholds for Endocrine Disrupters and related Uncertainties. Report of the Endocrine Disrupters Expert Advisory Group. JRC Scientific and Policy Report EUR 26068 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository	This report represents the opinions of The Endocrine Disrupters Expert Advisory Group (ED EAG) which is a sub-group of the ad hoc group of Commission Services, EU Agencies and Member States for the Community Strategy on Endocrine Disrupters. The report captures the experts' opinions on scientific issues relevant to the likelihood of the existence of thresholds for a biological response of an organism to an ED, in particular considering thresholds of adversity and the uncertainties associated with reliably estimating such thresholds from experimental data.

<p>OECD (2008). OECD Guideline for Testing of Chemicals: Test No 97: Detailed Review Paper on the Use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors. ENV/JM/MONO(2008)24, 95 pp. Available from: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2008)24&doclanguage=en</p>	<p>In this paper the need and importance to incorporate metabolising systems in <i>in vitro</i> methods for testing EDs is underlined. This notion is based on the fact that endogenous steroids and EDs are extensively metabolised by phase I and II enzymes and that such metabolism can lead to the activation or inactivation of steroids and, with regard to EDs to detoxification but also to the formation of active metabolites. This document justifies the incorporation of mammalian metabolising systems and indicates how this could be done with <i>in vitro</i> assays for EDs. The background to ED testing, the available test methods, and the role of mammalian metabolism in the activation and the inactivation of both endogenous and exogenous steroids is described. The available types of metabolising systems are compared, and the potential problems in incorporating metabolising systems in <i>in vitro</i> tests for EDs, and how these might be overcome, are discussed. Lastly, some recommendations for future activities are made.</p>
<p>OECD (2012). Series on Testing and Assessment: No 178: Detailed Review Paper on the State of the Science on Novel In vitro and In vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. ENV/JM/MONO(2012)23, 213 pp. Available from: http://www.oecd.org/env/ehs/testing/49002244.pdf</p>	<p>This paper describes endocrine pathways shown to be susceptible to environmental disruption and assays to address the relevant endpoints that could be used in new or existing OECD Test Guidelines for evaluating chemicals for endocrine-disrupting activity. In addition, the potential role of chemical-induced epigenetic modifications to endocrine signaling pathways, during sensitive windows of exposure, was evaluated as a mechanism of endocrine disruption, along with the examination of potential methods for assessing such disruption. Potential targets of disruption along putative adverse outcome pathways associated with the signaling pathways were identified, along with assays that show promise in evaluating the target in a screening and testing program. This document also proposes to the OECD Test Guidelines programme, suggestions of new assays or the incorporation of novel endpoints into existing assays that would expand the repertoire of endocrine signaling pathways included in the screening and testing regimen. This Detailed Review Paper was developed as a follow-up to the workshop on OECD countries' activities regarding testing, assessment and management of endocrine disruptors, which was held in Copenhagen in 2010.</p>
<p>WHO/IPCS (2002). Global Assessment of the State-of-the-science of Endocrine Disruptors. WHO/PCS/EDC/02.2, 180 pp.</p>	<p>In this document the authors aimed at evaluation of the entire body of knowledge concerning EDs coming from a vast number of research studies conducted under various conditions and examining various outcomes. Therefore it possess a unique feature for evaluating diverse data and it provides a framework and utilizes objective criteria for assessing causality between exposures to EDCs and selected outcomes.</p>
<p>WHO/UNEP (2013). State of the Science of Endocrine Disrupting Chemicals - 2012. 296 pp. Available from: http://www.who.int/mediacentre/news/releases/2013/hormone_disrupting_20130219/en/index.html</p>	<p>The United Nations Environment Programme (UNEP) and WHO presents an updated version of the IPCS (2002) document concerning the state of the science of endocrine disrupting chemicals. The document provides the global status of scientific knowledge on exposure to and effects of EDCs. It explains what endocrine disruption is all about, and then it discusses in detail, evidence of endocrine disruption in humans and wildlife.</p>

Zoeller RT, Brown TR, Doan LL, Gore AC, Skakkebaek NE, Soto AM, Woodruff TJ and Vom Saal FS, 2012. Endocrine-disrupting chemicals and public health protection: a statement of principles from The Endocrine Society. <i>Endocrinology</i> 153: 4097-4110.	This paper proposes the key considerations to be employed to identify EDCs and manage the risk to exposed populations as well as emphasizes the importance of the critical periods of development and the potential of low-dose EDC exposures to have potent and irreversible effects. Finally, recommendations for strengthening the EDC screening program through the incorporation of basic endocrine principles to promote further understanding of complex EDC effects are presented.
Mechanistic understanding of endocrine disruption in relation to non-monotonic dose response (reviews from the scientific literature)	
Beausoleil C, Ormsby J-N, Gies A, Hass U, Heindel JJ, Holmer ML, Nielsen PJ, Munn S and Schoenfelder G. 2013. Low Dose Effects and Non-monotonic Dose Responses for Endocrine Active Chemicals: Science to Practice Workshop. <i>Chemosphere</i> 93: 847 - 856	The paper summarises a workshop held in Berlin September 12–14th 2012 to assess the state of the science of the data supporting low dose effects and non-monotonic dose responses (“low dose hypothesis”) for chemicals with endocrine activity (endocrine disrupting chemicals or EDCs). This workshop consisted of lectures to present the current state of the science of EDC action and also the risk assessment process. These lectures were followed by breakout sessions to integrate scientists from various backgrounds to discuss in an open and unbiased manner the data supporting the “low dose hypothesis”. While no consensus was reached the robust discussions were helpful to inform both basic scientists and risk assessors on all the issues.
Borgert, C. J. Baker, S. P., Matthews, J.C. 2013. Potency matters: Thresholds govern endocrine activity. <i>Reg. Toxicol. Pharmac.</i> 67: 83 - 88	This brief review highlights how the fundamental principles governing hormonal effects – affinity, efficacy, potency, and mass action – dictate the existence of thresholds and why these principles also define the potential that exogenous chemicals might have to interfere with normal endocrine functioning. According to the authors vital signalling functions of the endocrine system require it to continuously discriminate the biological information conveyed by potent endogenous hormones from a more concentrated background of structurally similar, endogenous molecules with low hormonal potential. This obligatory ability to discriminate important hormonal signals from background noise can be used to define thresholds for induction of hormonal effects, without which normal physiological functions would be impossible. From such
EFSA Scientific Colloquium Summary Report, 14-15 June 2012. Low dose response in toxicology and risk assessment.	The low-dose effect and non-monotonicity hypotheses challenge key concepts in toxicology and risk assessment, and also the possibility to predict the effects of a chemical at low levels of

http://www.efsa.europa.eu/en/supporting/doc/353e.pdf	exposure from its effects at higher levels of exposure. This colloquium aimed to exchange views on the topics of low-dose effects and NMDRC and how these phenomena should
Rhomberg, L.R. & Goodman, J.E., 2012. Low-dose effects and non-monotonic dose-responses of endocrine disrupting substances: has the case been made. <i>Reg. Toxicol. Pharmacol.</i> 64: 130-133	The paper is a commentary critiquing Vandenberg et al. 2012 (see below) The authors state that the Vandenberg analysis framework lacks the scientific rigor necessary for an objective evaluation of the extent to which a body of scientific evidence does, or does not, support the NMDR hypothesis. In their view, the case for widespread non-monotonicity leading to undetected toxicity at low doses has not been made, and indeed cannot be made, simply through assembling selected cases that are presumed to represent causal effects
Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR, Jr., Lee DH, Shioda T, Soto AM, vom Saal FS, Welshons WV, Zoeller RT and Myers JP, 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. <i>Endocrine Reviews</i> 33: 378-455	In this paper two major concepts in EDC studies are discussed: low dose and non-monotonicity. The mechanistic data for low-dose effects and use of a weight-of-evidence approach to analyse examples from the EDC literature are reviewed. The authors illustrate that non-monotonic responses and low-dose effects are remarkably common in studies of natural hormones and EDCs. This paper concludes that when non-monotonic dose-response curves occur, the effects of low doses cannot be predicted by the effects observed at high doses. Thus, fundamental changes in chemical testing and safety determination are needed to protect human health.
<i>In vitro</i> methods	
Nielsen FK, Hansen CH, Fey J A, Hansen M, Jacobsen NW, Halling-Sørensen B, et al (2012). H295R cells as a model for steroidogenic disruption: a broader perspective using simultaneous chemical analysis of 7 key steroid hormones. <i>Toxicology in Vitro</i> 26: 343-350	In this paper the effects of three model endocrine disruptors, prochloraz, ketoconazole and genistein on steroidogenesis were tested in the adrenocortical H295R cell line. It was demonstrated that mechanistic differentiation was possible for these compounds in one assay by applying chemical analysis to the H295R assay. In general, the applied analytical method, analyzing 7 hormones simultaneously, enables the identification of important steps and differentiation of modes of action for the EDs investigated and the present findings correlate well with previous reports. The authors propose combining the present method with other techniques such as PCR and/or microarray to provide information on genes, enzymes and receptors involved in steroidogenesis and further aid to understand the mechanism by which EDs exert their effects on steroidogenesis. This could provide a better understanding of effects observed <i>in vivo</i> and mechanistic information or differentiation of modes of action.
Murk, A. J., Rijntjes, E., Blaauboer, B. J., Clewell, R., Crofton, K. M., Dingemans, M. M. L.Gutleb, A. C. (2013). Mechanism-based testing strategy using <i>in vitro</i> approaches for identification of thyroid hormone disrupting chemicals. <i>Toxicology in Vitro</i> 27: 1320–1346.	In this paper authors described and recommended a battery of test methods able to classify chemicals as of less or high concern for further hazard and risk assessment for TH disruption (THD). In addition, research gaps and needs were identified to optimize and validate the targeted THD <i>in vitro</i> test battery for a mechanism-based strategy for a decision to opt out or to proceed with further testing for THD.
Rotroff, D., Dix, D., Houck, K., Knudsen, T., MT, M., McLaurin, K. W., ... Judson, R. (2013). Using <i>in vitro</i> high throughput screening assays to identify potential endocrine-	This review analyses the briefly the history of the research and regulation in area of endocrine disruptors and proposes a new testing battery for high throughput screening for the endocrine disruptive potential of chemicals. The study uses U.S. EPA ToxCast HTS assays for estrogen,

<p>disrupting chemicals. Environmental Health Perspectives 121: 7–14.</p>	<p>androgen, steroidogenic, and thyroid-disrupting mechanisms to classify compounds and compare ToxCast results to <i>in vitro</i> and <i>in vivo</i> data from Endocrine Disruptor Screening Program (EDSP) T1S assays. The authors implemented an iterative model that optimized the ability of endocrine-related HTS assays to predict components of EDSP T1S and related results. The analysis shown that ToxCast estrogen receptor and androgen receptor assays predicted the results of relevant EDSP T1S assays with balanced accuracies and as well as Uterotrophic and Hershberger. Models for steroidogenic and thyroid-related effects could not be developed with the currently published ToxCast data. Overall, the results suggested that current ToxCast assays can accurately identify chemicals with potential to interact with the estrogenic and androgenic pathways, and therefore help prioritize chemicals for EDSP T1S assays.</p>
<p><i>In silico</i> methods</p>	
<p>Benfenati E, Roncaglioni A, Boriani E, Porcelli C, Spreafico M and Lo Piparo E, 2005. Validation of selected, non-commercial (Q)SAR models for Estrogen Receptor and Androgen Receptor binding. Final report of JRC contract CCR.IHCP.C430414.X0. 96 pp. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_Mario_Negri.pdf</p>	<p>A detailed literature review was conducted in order to identify all interesting publications related to ER and AR endpoints including bibliographic search, search through EC projects and the World Wide Web. Specific key points were collected, identifying suitable sources of data for testing that can serve as the basis for a new proposal for a high-quality QSAR. A rational scheme for applying the scores to each model was stated and its application permitted to identify most promising models at the top-level scores. Different kinds of model (regression and classification, SAR and QSAR models) have been analysed more in detail and when possible they were externally validated with data found in literature. Several ways for assessing the applicability domain have been evaluated depending on the available information. Currently available qualitative models (SAR and classification models) gave better results. Finally, based on the studies conducted during the project a rationale for developing a new high quality (Q)SAR is suggested for the future.</p>
<p>Lo Piparo E & Worth A (2010). Review of QSAR Models and Software Tools for Predicting Developmental and Reproductive Toxicity. JRC Scientific and Technical Report EUR 24522 EN Available from: http://publications.jrc.ec.europa.eu/repository</p>	<p>This JRC report provides a state-of-the-art review of available computational models for developmental and reproductive toxicity, including Quantitative Structure-Activity Relationship (QSARs) and related estimation methods such as decision tree approaches and expert systems. At the time of the publication, there were relatively few models for developmental and reproductive toxicity endpoints, and those available had limited applicability domains. This situation is partly due to the biological complexity of the endpoint, which covers many incompletely understood mechanisms of action, and partly due to the paucity and heterogeneity of high quality data suitable for model development. In contrast, there is an extensive and growing range of software and literature models for predicting endocrine-related activities, in particular models for oestrogen and androgen activity. There is a considerable need to further develop and characterise <i>in silico</i> models</p>

	for developmental and reproductive toxicity, and to explore their applicability in a regulatory setting.
OECD (2009b). Series on Testing and Assessment: No 111: Report of the Expert Consultation to Evaluate an Estrogen Receptor Binding Affinity Model for Hazard Identification. ENV/JM/MONO(2009)33, 119 pp.	This document is a report of an expert consultation held on 17 February 2009 with the aim to evaluate a QSAR approach for estimating estrogen receptor binding affinity for chemicals in defined regulatory inventories developed by the United States Environmental Protection Agency. The expert consultation was held based on the key recommendation from the OECD Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox held in May 2008 to develop structural alerts for identifying estrogen receptor binders for inclusion in the Toolbox during phase 2 development which started in November 2008.
Johnson, C. M., Achary, M., & Suri, R. P. (2013). An interaction model for estimating <i>in vitro</i> estrogenic and androgenic activity of chemical mixtures. Environmental Science & Technology 47: 4661-4669.	In this paper an alternate index, aRP, which enables the quantification of an antagonistic interaction from analytically derived concentrations of chemical constituents within a mixture that act upon the same molecular target is described. The index is calculated by measuring the degree to which the test compound modulates the activity of a standard hormone as a function of mixture proportions. The aRP was shown to be valid for additive mixtures. It theoretically estimates the product of the relative potential and the interaction index inverse for non-additive mixtures. The aRP values were computed for agonists and antagonists of both the estrogen and androgen receptors by using yeast-based methods (YES and YAS). The resulting aRP estimates were then validated using higher order mixtures of agonists and antagonists. The use of aRP led to improved predictions compared to estimates based on the toxicity equivalent factor (TEF) approach. The aRP model yielded estimates that were statistically indistinguishable ($\alpha = 0.01$) from the measured responses in 75% of the 32 mixtures tested. By the same criteria, the TEF approach successfully predicted 34% of the mixtures. Both the aRP and TEF approach correlated well with the observed responses (Pearson R = 0.98 and 0.84, respectively); however, the TEF estimates produced higher percent errors, particularly in mixtures with higher proportions of antagonists. It is suggested that the use of the aRP index allows for a better approximation of the net activity captured by the bioassays through the use of chemically derived concentrations.
Testing and assessment strategies	
Davis M, Boekelheide K, Boverhof DR, Eichenbaum G, Hartung T, Holsapple MP, Watkins PB (2013). The new revolution in toxicology: the good, the bad, and the ugly. Annals of the New York Academy of Sciences 1278: 11–24. http://onlinelibrary.wiley.com/doi/10.1111/nyas.12086/abstract;jsessionid=908D828E061387ACD8A0D582AAC66151.d02t01	This paper summarises the discussions held during the conference entitled “The New Revolution in Toxicology: The Good, Bad and Ugly” in 2011 hosted at the New York Academy of Sciences. The conference was focused on assessing the implementation of the vision presented in the report "Toxicity Testing in the 21st Century: A Vision and a Strategy", published by the United States National Academy of Sciences in 2007. The authors proposed the use some of the high-throughput assays from ToxCast to prioritize endocrine disruptor screening (EDSP21 program) as being the most advanced example ready for regulatory use. The article reviewed the achievements and the

	relevance of these achievements to safety assessment of new chemical entities intended for pharmaceutical use.
Kolle SN et al. (2012). A testing strategy for the identification of mammalian, systemic endocrine disruptors with particular focus on steroids. <i>Regulatory Toxicology and Pharmacology</i> : 63:259-278.	In this paper authors present an alternative testing paradigm for detection of endocrine modes of action that replace and reduce animal testing through refinement. The yeast-based receptor-mediated transcriptional activation YES/YAS assays were used to assess receptor mediated endocrine effects and effects on steroid hormone biosynthesis were assessed using the human cell line H295R in the steroidogenesis assay. The testing strategy was based on complementation of the in vitro assays with a single in vivo repeated dose study in which plasma samples are analyzed for their metabolome profile in addition to classical parameters such as histopathology. The authors claim that the combination of these methods does not only contribute to refinement and reduction of animal testing, but also has significantly increased the efficient allocation of resources and allows for a sound assessment of the endocrine disruption potential of compounds. Thus, this proposal constitutes a potentially attractive alternative to EPA's Endocrine Disruptor Screening Program to identify mammalian, systemic endocrine modes of action. Data on 14 reference substances for which the in vitro YES/YAS and steroidogenesis assays and the in vivo metabolome analysis were performed to assess their putative endocrine modes of action are presented.
Databases	
Ding D, Xu L, Fang H, Hong H, Perkins R, Harris S & Tong W (2010). The EDKB: an established knowledge base for endocrine disrupting chemicals. <i>BMC Bioinformatics</i> , 11 Suppl 6(Suppl 6), S5.	This project was initiated in the mid 1990's by the FDA as a resource for the study of EDs. The EDKB database, a component of the project, contains data across multiple assay types for chemicals across a broad structural diversity. The database contains 1,284 binding assays, reporter-gene assays, cell-proliferation assays, and in-vivo assays. It is cross-linked to other publicly available and related databases including TOXNET, Cactus, ChemIDplus, ChemACX, Chem Finder, and NCI DTP.
Shen J, Xu L, Fang H, Richard AM, Bray JD, Judson RS, & Hong H (2013). EADB: An Estrogenic Activity Database for Assessing Potential Endocrine Activity. <i>Toxicological Sciences</i> : 135(2), 277-91.	An Estrogenic Activity Database for Assessing Potential Endocrine Activity. EADB assembles a comprehensive set of estrogenic activity data from a variety of data sources and is a component of the enhanced Endocrine Disruptors Knowledge Base (EDKB). Searching can be carried out on structure (substructure search, super structure search, similarity search, full search, R-group search, and exclusion search) or on data, including numerical data (various estrogenic activity data) and text data (assay descriptions and literature references), as well as logical combinations of multiple searching operations.

12. Toxicokinetics

Varvara Gouliarmou, Sandra Coecke, Alicia Paini & Andrew Worth

12.1 In vivo sources of toxicokinetic information and regulatory guidelines

Toxicokinetics is the description of the penetration of a chemical into the body and its subsequent fate following exposure, which can for example be oral, dermal or through inhalation. Simply stated, toxicokinetics (TK) can be defined as what the body does to the chemical, as opposed to toxicodynamics (TD), which can be defined as what the chemical does to the body.

Traditionally, TK studies are performed in rats, according to OECD Test Guideline 417, by employing a minimum of four animals of each sex for each dose tested when no data are available. The measurement endpoints are the absorption rate, tissue distribution, metabolism and excretion of the chemical and its metabolites, together with TK parameters (bioavailability, Area under the curve [AUC], maximal plasma concentration [C_{\max}], time to reach C_{\max} [T_{\max}], clearance, half time). In the case of dermal exposure, *in vivo* studies are conducted according to the TG 427 and *in vitro* studies according to TG 428. EFSA has published a dermal absorption guidance document regarding the critical aspects related to the setting of dermal absorption values to be used in risk assessments of chemical plant protection products. EMA has published a guidance document (CPMP/ICH/384/95) regarding the TK assessment of pharmaceutical products intended for use in humans (Rowland, 2006).

The microdosing technique is another source of *in vivo* TK information (bioavailability and even metabolite identification) where human volunteers are used. In this technique, the TK of a drug is studied through the administration of doses lower than 1% of pharmacological active dose and up to a maximum of 100 μg . These doses are unlikely to produce whole-body effects, but are high enough to allow the cellular response to be studied. However, prior to microdosing, study of the likely risk of the chemical should be provided. This is achieved by a mixture of *in vitro* and animal (usually rat) data and in accordance to requirements laid down the latest ICH M3 (R2) guideline, which is now universally accepted (EMA, 2009).

The study of TK in mammals is a requirement for active ingredients in plant protection products (Regulation (EC) No 1107/2009) and in biocidal products (Regulation (EU) No 528/2012). In the REACH regulation, the conduct of new TK studies is not obligatory. However, Annex I (Section 1.0.2) requires the TK profile to be considered in the human health hazard assessment, and Annexes VIII-X require consideration of TK data if they are available.

12.2 Toxicokinetic processes - ADME

The four components of TK are Absorption, Distribution, Metabolism and Excretion, referred to collectively as ADME. The interplay between these processes defines the penetration and the fate of a compound in the body (Coecke et al., 2013).

Absorption is how a compound penetrates into the body. It is the process by which a compound passages from outer lining membranes such as the gut epithelium, the skin or the lung epithelium reaching subsequently the fluids and tissues/organs of the body. Compounds can be absorbed 1) passively through the cell membrane (transcellular route) or via the junctions between the cells (paracellular route) or 2) actively through carrier-mediation via an active or facilitated diffusion (Sarmiento et al., 2012). Additional, more than 400 types of efflux and influx transporters, localized in tissues and various cellular membranes of the human body, are functional and play a crucial role in absorption and disposition (Kathleen et al., 2010). The preferred pathway for absorption or transport of a specific compound depends

on its physicochemical characteristics and the biological features of the membrane (Sarmiento et al., 2012).

Distribution is the dispersion of a compound throughout fluids and tissues/organs of the body. This process describes how the compound is allocated to specific body compartments. After absorption, the distribution of a compound (and its metabolites) inside the body is governed by three main factors: 1) its partitioning between plasma fluid and plasma proteins; 2) its partitioning between blood and specific tissues; and 3) the ability of the substance to cross internal physiological interfaces, the so-called barriers (e.g. blood-brain barrier, blood-placenta barrier, blood-testis barrier) (Adler et al., 2011).

Metabolism (or biotransformation) is how the body chemically modifies a compound and it is the principal elimination route for most compounds. A vast number of xenobiotic metabolizing enzymes exist in human body (Klaassen, 2011). A toxic compound can be detoxified or a nontoxic compound may be transformed to toxic metabolite by biotransformation. Metabolism occurs in many tissues with liver being the most important organ, but also the kidneys, intestines, lungs, brain, nasal epithelium and skin can be involved (Asha & Vidyavathi, 2010). Generally, hepatic metabolism involves a phase I reaction where non polar molecules are converted into a polar compound and phase II conjugation where suitable moieties are added to the compound (Li, 2005).

Excretion is how the body irreversibly removes a chemical and its metabolites. Excretion occurs via kidneys (renal excretion), exhalation and less commonly via hepatobiliary system. Renal excretion is the results of three different mechanisms and they all include the interplay of both passive movement of the chemical and the participation of a number of active transporters. This process is mainly important for hydrophilic compounds. Exhalation is the primary excretion route for more volatile compounds. This process is mainly controlled by the affinity of compound between the blood and the air. Finally, biliary excretion is first preceded with the entry of the substance to the hepatocyte and its possible biotransformation by the phase I and/or phase II enzymes, with for some substances ultimately excreted into the bile (Adler et al., 2011).

12.3 Uses of toxicokinetic information

In the regulatory assessment of chemicals, TK information is useful in the design and interpretation of experimental studies, including animal experiments and non-animal testing strategies. A range of further possible applications is given in Table 12.1. In principle, all of these use cases could be addressed by using a suitable selection of non-standard methods. Some chemical-specific case studies illustrating how TK information can be used to improve experimental design and the final risk assessment are given by Bessems & Geraets (2013).

No ADME process occurs in isolation, instead complex interactions between the four processes results in a time dependent concentration of the compound in the systemic circulation or at a target site. In the vitro methods described below, cells and tissues are exposed to a fixed concentration. In contrast, in the body, cells and internal barriers are exposed to a time-dependent concentration profile of both parent compounds and formed metabolites as a result of the interplay of ADME processes. TK has been identified as a key element to integrate the results from *in silico*, *in vitro* and already available *in vivo* toxicity studies.

TK is needed to estimate the range of target organ doses that can be expected from realistic external exposure scenarios. This information is crucial for determining the dose range that should be used for *in vitro* testing. Vice versa, TK is necessary to convert the *in vitro* results, generated at tissue/cell or sub-cellular level, into dose response or potency information relating to the entire target organism, i.e. the human body (*in vitro* – *in vivo* extrapolation).

Physiologically based toxicokinetic modelling (PBTK) is currently regarded as the most adequate approach to simulate human TK. The fact that PBTK models are mechanism-based and therefore allows them to be 'generic' to a certain extent (various extrapolations possible) has been critical for their success so far. The need for high-quality *in vitro* and *in silico* TK data as input for PBTK models to predict human dose-response curves is currently a bottleneck for integrative risk assessment. The full replacement of current animal toxicokinetics tests by PBTK modelling, based on *in silico* and *in vitro* data, will take many years of increased efforts.

TK information is crucial in read-across and grouping of chemicals. Verification that similar TK properties apply to the members of a suggested group/category of chemicals is considered to be one of the steps in building a case of a plausible read-across case under the REACH Regulation. This is acknowledged in Annex XI, 1.5 according to which similarities among the group members may be based e.g. on "the common precursors and/or the **likelihood of common break-down products via physical and biological processes**, which result in structurally similar chemicals".

Under REACH there is a possibility to waive/adapt certain information requirements. The sub-chronic study can be waived, if "the substance is unreactive, insoluble and not inhalable and there is **no evidence of absorption** and no evidence of toxicity in a 28-day 'limit test', particularly if such a pattern is coupled with limited human exposure." (REACH Regulation Annex IX, 8.6.2. Column 2). Furthermore, pre-natal developmental toxicity and two generation reproductive toxicity test can be waived, if "the substance is of low toxicological activity (no evidence of toxicity seen in any of the tests available), **it can be proven from toxicokinetic data that no systemic absorption occurs via relevant routes of exposure (e.g. plasma/blood concentrations below detection limit using a sensitive method and absence of the substance and of metabolites of the substance in urine, bile or exhaled air)** and there is no or no significant human" (Annex IX, 8.7. Column 2). Obviously, generation of toxicokinetic data is a prerequisite, before these alternative approaches can be used.

It is noteworthy that since there is no *in vivo* information requirement under REACH Regulation, the choice of scientifically valid *in vitro* methods does not need to be justified, when *in vitro* TK data is generated for specific regulatory purposes, for example those summarised above.

12.4 Non-standard methods for ADME

A large and rapidly growing body of scientific literature refers to the development of non-standard methods (*in vitro* and *in silico*) for describing and predicting ADME processes (see Table of References with Notes for recent reviews). These include a variety of *in vitro* methods, the most widely used of which are presented in Table 12.2 together with the measured parameters that can be used as input data for PBTK models.

12.4.1 *In vitro* methods for absorption through external barriers

For the prediction of *in vivo* absorption, *in vitro* permeability studies are performed. Permeability methods include two diffusional compartments that are separated by cell lines or primary cells or tissues of human/animal origin, or by artificial/biological membranes. The choice of the barrier between the two compartments depends on the exposure route, while the ability of the model to predict chemical's absorption depends on how closely it mimics the characteristics of the *in vivo* situation. Cell lines are often employed as a barrier and their ability to form tight junctions is critical for the performance of the *in vitro* assay. Various cell lines like Caco-2, TC-7, HT29-MTX and IEC-18 have been employed to predict intestinal absorption (Grès et al., 1998; Behrens et al., 2001; Versantvoorta et al., 2002). Caco-2 cell line is the foremost used and well characterized and it has been widely used to generate drug

absorption data for regulatory purposes. Caco-2 is a very good model to mimic passive transport through barriers and it is believed that transporter proteins are functional on its membranes. A disadvantage of Caco-2 model stems from the wide variation with passage number and differing results obtained across different laboratories (Zucco et al., 2005; Prieto et al., 2010; reviewed in Nigsch F et al., 2007). The parallel artificial membrane permeability assay (PAMPA) is also a frequently used system to assess intestinal permeability (Kansy et al., 1998). PAMPA model only deals with exclusive passive diffusion, while some intestinal cell lines attempt to capture some aspects of active transport across the intestinal barrier (Nigsch F et al., 2007). Absorption through the respiratory tract can be predicted using the respective cell lines or primary cell cultures. Concerning the human bronchial epithelia cell lines, Calu-3 (Zhu, 2010) and 16HBE14o- (Forbes et al., 2003) cell monolayers have been used as models for airway epithelium due to their morphological characteristics, barrier properties and expression of drug transport systems that exist in vivo (Zhu, 2010). Regarding the alveolar region, A549 cell lines have been employed either as monolayers (Wang & Zhang, 2004) or in co-culture with blood monocyte-derived macrophages and dendritic cells (Lehmann et al., 2011). There are also commercially available 3D systems based on human derived tracheal/bronchial epithelia cells (Chemuturi et al., 2005). For nasal permeation studies, monolayers obtained with primary cultures of human epithelial cells have been extensively used (Kissel & Werner, 1998; Agu et al., 2002; Agu et al., 2011; Sarmiento et al., 2012). Additionally, cell lines such as RPMI 2650 (human), BT (bovine) and NAS 2BL (rat) have been employed in nasal drug development (Dimova et al., 2005), while recently co-cultures composed by a collagen matrix with embedded human nasal fibroblasts covered by a RPMI 2650 were developed (Wengst & Reichl, 2010). Skin absorption can be predicted by employing reconstructed skin models that simulate human skin (Godin & Touitou, 2007). Various cultured human skin equivalents such as living skin equivalent models (LSEs) and the full thickness skin models have been used to measure percutaneous absorption, with some of them having also metabolic competence (Jäckh C et al., 2011). Commercially available in vitro human 3D skin models like Episkin, EpiDerm and Skin Ethic, employ epidermal cells (Netzlaff et al., 2005), the skin-PAMPA model uses an artificial membrane based on phospholipids, cholesterol or porcine polar dissolved in a solvent (Sinkó et al., 2012), these commercial models have been the most developed and used. The 3D reconstructed human skin models are able to mimic human skin to a large extent compared to the classical cell monolayers. Also real skin tissues have been implemented as membrane in ex vivo tests. According to the OECD 428 guideline excised skin can be used in in vitro methods, with human skin being the most acceptable followed by animal skin. Although rodent have been extensively used in in vivo experiments (OECD 417) it is well documented that absorption through rodent skin is greater than through human skin due to the potential for compounds to obtain rapid passage down the hair follicle and the relative thickness of the stratum corneum (Williams, 2006). However, pig ear skin is a more relevant animal tissue for ex vivo studies due to its more structural equivalence to human skin (Godin & Touitou, 2007).

12.4.2 In vitro methods for distribution

The prediction of a chemical's distribution within the body requires measurements of the free (unbound) compound fraction in blood, the distribution coefficients and permeability through internal physiological interfaces. Partitioning methods to determine free fractions and distribution coefficients were reviewed by Heringa and Hermens (Heringa & Hermens, 2003). The ultracentrifugation method employs powerful centrifugation (e.g. 250,000 g) and long centrifugation times (e.g. 16 hours) to separate the free from the bound form of the compound. The aim is to precipitate the binding matrix and then to measure the free compound concentration in the supernatant. In ultrafiltration method the phase separation between the bound and the free form is achieved by filtration using a special size-excluding filter. Filtration is performed either by applying pressure or by centrifugation and the free concentration is measured in the filtrate. Rapid equilibrium dialysis (RED) method employs a

device consisting of two compartments separated by a membrane. The one compartment contains the sample and the other compartment a buffer solution. Only the free form of the compound can cross the membrane. When equilibrium between the two compartments is established, the freely dissolved concentration of the two compartments is the same. The concentration of the compound in the matrix free compartment is measured and equals with the free concentration. Solid Phase Microextraction (SPME) is a method that requires no phase separation step and thus more precise since the phase separation can shift the equilibrium between the free and the bound form. SPME employs a small amount of a polymer as a third partitioning phase. The polymer is immobilized on a solid support (usually a fibre) and then is brought in equilibrium with the sample that contains both the free and the bound form. Only a minor mass of the free form of the compound can be absorbed/adsorbed by the polymer and it is consequently extracted from the polymer and quantified. The head space sampling method is useful to determine the free concentration of volatile compounds. The sample is placed inside a closed vial and is then equilibrated with the headspace. The measured concentration of the compound in the headspace is proportional to its free concentration in the sample. Usually the sampling of the head space is performed by inserting and equilibrating a SPME fibre, in this case the method is called head space SPME.

The distribution process is also influenced by the passage through internal barriers such as blood-brain barrier, blood-testis barrier and placenta. The blood-brain barrier is one of the most important internal interfaces within the human body. Various *in vitro* methods have been developed during the last decades aiming to determine the permeability of chemicals through internal barriers. Naim and Cucullo have recently reviewed blood-brain barrier technologies (Naik & Cucullo, 2012). It seems that the *in vitro* cell based models are the most commonly used, since they can be established with any type of cell source (human, animal or cell line derivative) including endothelial cells and freshly isolated astrocytes. The primary bovine brain endothelial cells are employed and they are probably the most suitable *in vitro* model (Culot M et al., 2008; Gumbleton & Audus, 2001). Due to their limited availability and the difficulty to established and maintain primary cultures Garberg et al tried to identify an alternative cell system by comparing different cell lines (Garberg et al., 2005). They observed that higher correlations with *in vivo* studies were seen when only passively transported compounds were included in the analysis of the *in vitro* results. However no linear correlation between *in vivo* and *in vitro* permeability was found for any of the *in vitro* models when the whole set of tested compounds, reflecting different transport mechanism and different degrees of permeability, was taken into account.

12.4.3 In vitro methods for metabolism

In vitro assessment of metabolism is achieved using tissue preparations, whole cells, primary cell lines or cell extracts. In these systems the enzyme conditions can be optimized and controlled in order to measure catalytic processes, such as phase I and phase II biotransformation, to isolate bioactivation from detoxification (Lipscomb & Poet, 2008) and to enable inter/intra species comparison (Walker SA et al., 2006). Metabolism occurs to the higher extend in the liver although other organs/tissues (lungs, small intestine and skin) are contributing to the overall metabolisms. Primary cultures of human hepatocytes are a very relevant *in vitro* liver metabolism system (Billings et al., 1977) and are considered for many applications as the human metabolic competent *in vitro* standard. However, obtaining freshly isolated hepatocytes from human donors has practical difficulties. As a result cryopreserved human hepatocytes have become a viable alternative since they can maintain more than 90% of fresh hepatocyte activity for at least a year after preparation (Richert et al., 2010; McGinnity et al., 2004.). Metabolic competent system, such as human cryopreserved hepatocytes and the HepaRG cell line, have shown good data on metabolic clearance, cytochrome P450 (CYP) induction and metabolite identification (Gerin et al., 2013; Jouin et al., 2006).

The HepaRG cell line is an important *in vitro* tool and a good alternative to human hepatocytes. This cell line originates from the liver of a human female donor and it exhibits a long-term stability of biotransformation functions (Andersson et al., 2012). However, since HepaRG cells are delivered from a single donor, they do not represent the whole spectrum of inter-individual variances in the human liver as compared to human hepatocytes. HepaRG have been shown to provide reliable prediction and characterization of CYP induction by drug compounds in humans as well as correct plasma membrane transporter polarization (Andersson et al., 2012). In view of these properties, HepaRG cells are suitable for use in two-dimensional (2D) and three-dimensional (3D) models, such as liver bioreactors (Leite et al., 2012) which enable studies of disposition and toxicity involving uptake, metabolism and efflux mechanisms over an extended period of time, since the expression of CYP enzymes is maintained in the differentiated HepaRG cells for at least 4 weeks (Jossé R et al., 2008).

During recent years, attempts have focused on producing metabolically active hepatocytes from human embryonic stem (ES) cells and hepatic stem cells, with several reports describing the development and optimization of various differentiation protocols. However, the applicability of the human ES cell-derived hepatocyte-like cells as a reliable tool to predict metabolism is under question since studies have shown that the key CYP3A4 isoform is present at 1000-fold lower amounts in differentiated ES cells than in human primary hepatocytes (Wobus et al., 2011). Recent progress has also been made in generating hepatocyte-like cells from human induced pluripotent stem (PS) cells. A direct comparison of differentiated ES cells and induced PS cells with human primary hepatocytes has shown that the liver specific functions (glycogen synthesis, urea production, albumin secretion and CYP activity) of the stem-cell derived cells were present but much lower compared to human hepatocytes (Song et al 2009). At present it is not clear to what degree stem cell-derived hepatocytes need to resemble primary hepatocytes to ensure a reliable metabolically competent system and critical issues must be resolved before large numbers of metabolic competent hepatocytes can be generated from such cell lines and used for ADME testing (Pelkonen et al., 2013).

Subcellular fractions, like microsomal protein (MSP), cytosol and S9 fractions have been used for decades in studies of xenobiotic metabolism. An advantage of these preparations is that they are relative stable over time and their metabolic activity is often initiated by the addition of substrate or cofactor, which practically means that the incubation time is easily controlled (Lipscomb & Poet, 2008). A disadvantage is that these preparations maintain neither the balance of phase I to phase II enzymes nor the ratio of cytosolic versus membrane bound enzymes (Lipscomb & Poet, 2008). Furthermore, they lack a cellular membrane barrier and maintain higher levels of metabolizing enzyme activity than hepatocytes (Alqahtani et al., 2013). Tissue slices have been employed in *in vitro* methods on the assumption that they provide a more biologically relevant system; however they do not come without limitations. One limitation is that the slice preparation damages an important cell proportion and can make them metabolically incompetent. Also the slow diffusion of the compound within the tissue may lead to underestimation of the metabolic rates (Lipscomb & Poet, 2008).

12.4.4 In vitro methods for elimination

In vitro renal excretion tools for humans have not been developed, but this process can be indirectly determined using physiologically parameters. The renal excretion rate can be indirectly predicted by the combination of the glomerular filtration rate (GRF) and the free fraction of the compound in the blood, while excretion via exhalation can be predicted using blood:air distribution coefficients (Bessemers et al., 2014). Cryopreserved hepatocytes cultured in a sandwich configuration allow for the development of intact bile canaliculi and the ability to measure hepatic uptake and biliary clearance (Bi Y-a et al., 2006).

12.4.5 *In silico* modelling

In silico modelling covers a wide range of approaches, and these have been extensively applied to the prediction of ADME properties, as reviewed by Mostrag-Szlichtyng & Worth (2010). The main approaches include simple rules-of-thumb based on structural alerts and physicochemical properties (e.g. the Lipinski rule-of-five), QSAR models, expert systems for simulating metabolism, models of receptor-mediated interactions (e.g. cytochromes), and physiologically based toxicokinetic (PBTK) models.

A range of QSAR software tools include the capacity for predicting key ADME properties, such as human intestinal absorption, oral bioavailability, skin penetration, blood/brain barrier permeability, plasma protein binding, and metabolic clearance. Many of these models have been developed by using data sets skewed towards drugs. These models are easy to apply, but their applicability to non-pharmaceutical chemicals is not clear.

Several expert systems for simulating metabolic pathways are also available, mainly in the commercial domain. These models generally provide qualitative information on possible metabolic pathways (with a tendency to overgenerate metabolites), but are not yet capable of predicting rates of metabolite formation.

PBTK models are mechanistically-based mathematical models that integrate physicochemical and biological data in order to simulate the dose and time-dependent concentration profiles of a chemical through the systems of the body. Different compartments in the model correspond to different organs. The transfer between compartments is described by ordinary differential equations and these are connected with hypothetical blood flows mimicking the blood circulation in the body. Organ sizes, blood flow rates and tissue-plasma distribution coefficients are used to construct the model, which is based on mass balance assumptions. The differential equations are solved numerically with mathematical software packages. There is no universally suitable model - the complexity of the model can be tailored to the needs of the study. For instance, for the simple screening of chemicals it may be sufficient to take into account only passive processes between a limited number of compartments (e.g. Tonnelier et al, 2012). PBTK models also provide a means of extrapolating between doses and species, and of accounting for inter-individual variation. Compared with other *in silico* approaches, these models are relatively data-hungry and need to be carefully optimised. Although in the past, these models required specialised expertise to develop and apply, intense efforts are undertaken to strive for public available models that can be used for routine regulatory applications and thus simplified approaches are proposed (Bessemers et al., 2014).

12.4.6 *Body-on-chip* methods

During recent years the development of microfabrication technology and its combination with cell culture techniques led to the emergence of microfluid devices that attempt to reproduce the multi-organ interactions. A microfluid device is fabricated with multiple chambers. Each chamber represents a different organ and all are connected with fluid conduits representing the blood flow; this new technology is called "body-on-a-chip" and provides the possibility to build up physical systems that mimics PBTK models (Sung et al., 2010). The concept behind the development of microfluid devices is promising; however these devices still pose some limitations (Kim et al., 2007) and many issues need to be addressed before they can be standardized. For instance, it is documented that the biochemical properties of cells differ in macro and microcultures, thus proper cell culture protocols suitable for these miniaturized platforms should be developed and employed (Su et al., 2013).

12.5 Conclusions

TK information provides a useful supporting role in the regulatory assessment of chemicals, and does not necessarily need to be generated by traditional animal methods. There are multiple applications of TK methods, ranging from the relatively simple (e.g. supporting read-across of toxicity data between analogues) to the complex (e.g. enabling animal-free risk assessments). To support these applications, a wide and diverse range of non-standard TK methods have been developed, including both *in vitro* and *in silico* tools, and mathematical modelling can be used to integrate multiple TK properties to simulate the dose and time-dependent disposition of a chemical in the body. The relevance of TK properties is self-evident, but the usefulness of individual TK methods can only be judged in the context of the intended application. In this respect, there is a need to develop further guidance on the use of TK information, and to develop standards for comparing the characteristics of new methods with more established ones.

12.7 References

- Adler S, Basketter D, Creton S, Pelkonen O, Van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleijnans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, Van Delft J, Van Loveren H, Vinken M, Worth A and Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: Current status and future prospects-2010. *Archives of Toxicology* 85(5): 367-485.
- Agu RC, MacDonald E, Cowley D, Shao K, Clarke DB & Massoud E (2011). Differential expression of organic cation transporters in normal and polyps human nasal epithelium: Implications for in vitro drug delivery studies. *International Journal of Pharmaceutics* 406(1-2): 49-54.
- Agu RU, Jorissen M, Kinget R, Verbeke N & Augustijns P (2002). Alternatives to in vivo nasal toxicological screening for nasally-administered drugs. *S.T.P. Pharma Sciences* 12(1): 13-22.
- Alqahtani S., Mochamed A. L., Kaddoumi A. (2013). Experimental models for predicting drug absorption and metabolism. *Expert Opinion on Drug Metabolism & Toxicology* 9(10): 1-14.
- Andersson TB, Kanebratt KP & Kenna JG (2012). The HepaRG cell line: A unique in vitro tool for understanding drug metabolism and toxicology in human. *Expert Opinion on Drug Metabolism and Toxicology* 8(7): 909-920.
- Asha S & Vidyavathi M (2010). Role of human liver microsomes in in vitro metabolism of drugs-A review. *Applied Biochemistry and Biotechnology* 160(6): 1699-1722.
- Behrens I, Stenberg P, Artursson P & Kissel T. (2001). Transport of Lipophilic Drug Molecules in a New Mucus-Secreting Cell Culture Model Based on HT29-MTX Cells. *Pharmaceutical Research* 18(8): 1138-1145.
- Bessems JGM & Geraets L (2013). Proper knowledge on toxicokinetics improves human hazard testing and subsequent health risk characterisation. A case study approach. *Regulatory Toxicology and Pharmacology* 67: 325-334.
- Bessems JG, Loizou G, Krishnan K, Clewell HJI, Bernasconi C, Bois F, Coecke S, Collnot EM, Diembeck, WLR Farcal, Geraets L, Gundert-Remy U, Kramer N, Küsters G, Leite SB, Pelkonen OR, Schröder K, Testai E, Wilk-Zasadna I & Zaldivar-Comenges J.M. (2014). PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment. Recommendations from a joint EURL ECVAM - EPAA workshop. *Regulatory Toxicology and Pharmacology* 68: 119-139.
- Billings RE, McMahon RE, Ashmore J and Wagle SR (1977). The metabolism of drugs in isolated rat hepatocytes. A comparison with in vivo drug metabolism and drug metabolism in subcellular liver fractions. *Drug Metabolism and Disposition* 5(6): 518-526.
- Bi Y.-a, Kazolias D & Duignan DB (2006). Use of Cryopreserved Human Hepatocytes in Sandwich Culture to Measure Hepatobiliary Transport. *Drug Metabolism and Disposition* 34(9): 1658-1665.
- Chemuturi NV, Hayden P, Klausner M & Donovan MD (2005). Comparison of human tracheal/bronchial epithelial cell culture and bovine nasal respiratory explants for nasal drug transport studies. *Journal of Pharmaceutical Sciences* 94(9): 1976-1985.
- Coecke S, Pelkonen O, Leite SB, Bernauer U, Bessems JGM, Bois FY, Gundert-Remy U, Loizou G, Testai E & Zaldivar JM (2013). Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicology in Vitro* 27(5): 1570-1577.

- Culot M, Lundquist S, Vanuxeem D, Nion S, Landry C, Delplace Y, Dehouck MP, Berezowski V, Fenart L, Cecchelli R. (2008). An in vitro blood-brain barrier model for high throughput (HTS) toxicological screening. *Toxicology in Vitro* 22(3): 799-811.
- Dimova S, Brewster ME, Noppe M, Jorissen M and Augustijns P (2005). The use of human nasal in vitro cell systems during drug discovery and development. *Toxicology in Vitro* 19(1): 107-122.
- EMA (2009). ICH guideline M3(R2) on non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals.
- Forbes B, Shah A, Martin GP and Lansley AB (2003). The human bronchial epithelial cell line 16HBE14o- as a model system of the airways for studying drug transport. *International Journal of Pharmaceutics* 257(1-2): 161-167.
- Garberg P, Ball M, Borg N, Cecchelli R, Fenart L, Hurst RD, Lindmark T, Mabondzo A, Nilsson JE, Raub TJ, Stanimirovic D, Terasaki T, Öberg JO & Österberg T (2005). In vitro models for the blood-brain barrier. *Toxicology in Vitro* 19(3): 299-334.
- Godin B & Touitou E (2007). Transdermal skin delivery: Predictions for humans from in vivo, ex vivo and animal models. *Advanced Drug Delivery Reviews* 59(11): 1152-1161.
- Grès MC, Julian B, Bourrié M, Meunier V, Roques C, Berger M, Boulenc X, Berger Y & Fabre G (1998). Correlation Between Oral Drug Absorption in Humans, and Apparent Drug Permeability in TC-7 Cells, A Human Epithelial Intestinal Cell Line: Comparison with the Parental Caco-2 Cell Line. *Pharmaceutical Research* 15(5): 726-733.
- Gumbleton M, & Audus KL (2001). Progress and limitations in the use of in vitro cell cultures to serve as a permeability screen for the blood-brain barrier. *Journal of Pharmaceutical Sciences* 90(11): 1681-1698.
- Heringa MB & Hermens JLM (2003). Measurement of free concentrations using negligible depletion-solid phase microextraction (nd-SPME). *Trends in Analytical Chemistry* 22(9): 575-587.
- Jäckh C, Blatz V, Fabian E, Guth K, van Ravenzwaay B, Reisinger K, Landsiedel R (2011). Characterization of enzyme activities of Cytochrome P450 enzymes, Flavin-dependent monooxygenases, N acetyltransferases and UDP-glucuronyltransferases in human reconstructed epidermis and full thickness skin models. *Toxicology In Vitro* 25(6):1209-14.
- Jossé R, Aninat C, Glaise D, Dumont J, Fessard V, Morel F, Poul JM, Guguen-Guillouzo C & Guillouzo A (2008). Long-Term Functional Stability of Human HepaRG Hepatocytes and Use for Chronic Toxicity and Genotoxicity Studies. *Drug Metabolism and Disposition* 36(6): 1111-1118.
- Jouin D, Blanchard N, Alexandre E, Delobel F, David-Pierson P, Lavé T, Jaeck D, Richert L & Coassolo P (2006). Cryopreserved human hepatocytes in suspension are a convenient high throughput tool for the prediction of metabolic clearance. *European Journal of Pharmaceutics and Biopharmaceutics* 63(3): 347-355.
- Kansy M, Senner F & Gubernator K (1998). 1. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *Journal of Medicinal Chemistry* 41(7): 1007-1010.
- Kathleen M, Giacomini SMH, Tweedie DJ, Benet LZ, Brouwer KLR, Chu X, Dahlin A, Evers R, Fischer V, Hillgren KM, Hoffmaster KA, Ishikawa T, Keppler D, Kim RB, Lee CA, Niemi M, Polli JW, Sugiyama Y, Swaan PW, Ware JA, Wright SH, Yee SW, Zamek-Gliszczynski MJ & Zhang L (2010). Membrane transporters in drug development. *Nat Rev Drug Discov* 9(3): 215-236.
- Kim L, Toh YC, Voldman J & Yu H (2007). A practical guide to microfluidic perfusion culture of adherent mammalian cells. *Lab on a Chip* 7(6): 681-694.

Kissel T. & Werner U. (1998). Nasal delivery of peptides: an in vitro cell culture model for the investigation of transport and metabolism in human nasal epithelium. *Journal of Controlled Release* 53(1–3): 195-203.

Klaassen CD (2011). *Casarett & Doull's Toxicology: The Basic Science of Poisons* (6th Edition) McGraw-Hill, New York, NY.

Lehmann AD, Daum N, Bur M, Lehr CM, Gehr P & Rothen-Rutishauser BM (2011). An in vitro triple cell co-culture model with primary cells mimicking the human alveolar epithelial barrier. *European Journal of Pharmaceutics and Biopharmaceutics* 77(3): 398-406.

Leite SB, Wilk-Zasadna I, Zaldivar JM, Airola E, Reis-Fernandes MA, Mennecozi M, Guguen-Guillouzo C, Chesne C, Guillou C, Alves PM and Coecke S (2012). Three-Dimensional HepaRG Model as an Attractive Tool for Toxicity Testing. *Toxicological Sciences* 130(1): 106-116.

Li AP, (2005). Preclinical in vitro screening assays for drug-like properties. *Drug Discovery Today: Technologies* 2(2): 179-185.

Lipscomb JC & Poet TS (2008). In vitro measurements of metabolism for application in pharmacokinetic modeling. *Pharmacology & Therapeutics* 118(1): 82-103.

McGinnity DF, Soars MG, Urbanowicz RA & Riley RJ (2004). Evaluation of fresh and cryopreserved hepatocytes as in vitro drug metabolism tools for the prediction of metabolic clearance. *Drug Metabolism and Disposition* 32(11): 1247-1253.

Mostrag-Szlichtyng Aleksandra & Worth Andrew (2010). Review of QSAR models and software tools for predicting biokinetic properties. European Commission. Available from: <http://publications.jrc.ec.europa.eu/repository>

Naik P & Cucullo L (2012). In vitro blood–brain barrier models: Current and perspective technologies. *Journal of Pharmaceutical Sciences* 101(4): 1337-1354.

Netzlaff F, Lehr CM, Wertz PW & Schaefer UF (2005). The human epidermis models EpiSkin®, SkinEthic® and EpiDerm®: An evaluation of morphology and their suitability for testing phototoxicity, irritancy, corrosivity, and substance transport. *European Journal of Pharmaceutics and Biopharmaceutics* 60(2): 167-178.

Nigsch F, Klaffke W, Miret S, (2007). In vitro models for processes involved in intestinal absorption. *Expert Opinion on Drug Metabolism & Toxicology* 3(4): 545-556.

Pelkonen O, Turpeinen M, Hakkola J, Abass K, Pasanan M, Raunio H & Vähäkangas K (2013). Preservation, induction or incorporation of metabolism into the in vitro cellular system- Views to current opportunities and limitations. *Toxicology in Vitro* 27: 1578-1583.

Prieto P, Hoffmann S, Tirelli V, Tancredi F, González I, Bermejo M, De Angelis I (2010). An exploratory study of two Caco-2 cell models for oral absorption: a report on their within-laboratory and between-laboratory variability, and their predictive capacity. *Altern Lab Anim.* 38(5):367-86.

Richert L, Abadie C, Bonet A, Heyd B, Manton G, Alexandre E, Bachellier P, Kingston S, Pattenden C, Illouz S, Dennison A, Hoffmann S & Coecke S (2010). Inter-laboratory evaluation of the response of primary human hepatocyte cultures to model CYP inducers – A European Centre for Validation of Alternative Methods (ECVAM) – Funded pre-validation study. *Toxicology in Vitro* 24(1): 335-345.

Rowland, M. (2006). *Microdosing and the 3Rs*. National Center for the Replacement, Refinement and Reduction of Animals in Research.

Sarmento B, Andrade F, Da Silva SB, Rodrigues F, Das Neves J & Ferreira D (2012). Cell-based in vitro models for predicting drug permeability. *Expert Opinion on Drug Metabolism and Toxicology* 8(5): 607-621.

Sinkó B, Garrigues TM, Balogh GT, Nagy ZK, Tsinman O, Avdeef A & Takács-Novák K. (2012). Skin-PAMPA: A new method for fast prediction of skin penetration. *European Journal of Pharmaceutical Sciences* 45(5): 698-707.

Song Z, Cai J, Liu Y, Zhao D, Yong J, Duo S, Song X, Guo Y, Zhao Y, Qin H, Yin X, Wu C, Che J, Lu S, Ding M & Deng H. (2009). Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. *Cell Research* 19(11): 1233-1242.

Su X, Theberge AB, January CT and Beebe DJ (2013). Effect of Microculture on Cell Metabolism and Biochemistry: Do Cells Get Stressed in Microchannels? *Analytical Chemistry* 85(3): 1562-1570.

Sung JH, Esch MB & Shuler ML (2010). Integration of in silico and in vitro platforms for pharmacokinetic/pharmacodynamic modeling. *Expert Opinion on Drug Metabolism and Toxicology* 6(9): 1063-1081.

Swift B, Pfeifer ND, Brouwer KLR. (2013). Sandwich-cultured hepatocytes: an in vitro model to evaluate hepatobiliary transporter-based drug interactions and hepatotoxicity. *Drug Metabolism Reviews* 42(3): 446-471.

Tonnelier A, Coecke S & Zaldivar JM (2012). Screening of chemicals for human bioaccumulative potential with a physiologically based toxicokinetic model. *Archives of Toxicology* 86: 393-403.

Versantvoorta CHM, Ondrewater RCA, Duizera E, Van de Sandtb JJM, Gildeb AJ, Grotenb JP (2002). Monolayers of IEC-18 cells as an in vitro model for screening the passive transcellular and paracellular transport across the intestinal barrier: comparison of active and passive transport with the human colon carcinoma Caco-2 cell line. *Environmental Toxicology and Pharmacology* 11(3-4): 335-344.

Walker SA, Whitten LB, Seals GB, Lee WE, Archibong AE & Ramesh A (2006). Inter-species comparison of liver and small intestinal microsomal metabolism of fluoranthene. *Food and Chemical Toxicology* 44(3): 380-387.

Wang Z & Zhang Q (2004). Transport of proteins and peptides across human cultured alveolar A549 cell monolayer. *International Journal of Pharmaceutics* 269(2): 451-456.

Wengst A & Reichl S (2010). RPMI 2650 epithelial model and three-dimensional reconstructed human nasal mucosa as in vitro models for nasal permeation studies. *European Journal of Pharmaceutics and Biopharmaceutics* 74(2): 290-297.

Williams FM (2006). In vitro studies—how good are they at replacing in vivo studies for measurement of skin absorption? *Environmental Toxicology and Pharmacology* 21(2): 199-203.

Wobus AM & Löser P (2011). Present state and future perspectives of using pluripotent stem cells in toxicology research. *Archives of Toxicology* 85: 79-117.

Zhu Y, Chidekel A & Shaffer TH (2010). Cultured human airway epithelial cells (calu-3): a model of human respiratory function, structure, and inflammatory responses. *Critical Care Research and Practice*. Article ID 394578. <http://www.hindawi.com/journals/ccrp/2010/394578>

Zucco F, Batto AF, Bises G, Chambaz J, Chiusolo A, Consalvo R, Cross H, Dal Negro G, de Angelis I, Fabre G, Guillou F, Hoffman S, Laplanche L, Morel E, Pinçon-Raymond M, Prieto P, Turco L, Ranaldi G, Rousset M, Sambuy Y, Scarino ML, Torreilles F & Stamatii A (2005). An inter-laboratory study to evaluate the effects of medium composition on the differentiation and barrier function of Caco-2 cell lines. *Altern Lab Anim* 33(6):603-18.

Table 12.1. Applications of toxicokinetic methods

Use Case	Examples
1) Filling data gaps by extrapolating between species, exposure routes, and exposure durations.	<p>1) In developmental toxicity testing, rodent data are available. Is there a need for non-rodent data?</p> <p>2) How can I extrapolate from oral repeat dose toxicity data to assess the effects of exposure via the dermal route?</p> <p>3) How can I extrapolate from a subchronic (90d) repeat-dose toxicity study to a long-term study?</p>
2) Optimal design of in vitro experiments for in vitro-in vivo extrapolation and reverse dosimetry. Animal-free risk assessment.	For a given exposure scenario (e.g. dietary exposure) how do I design an in vitro toxicity study suitable for assessing repeat-dose toxicity?
3) Optimal design of animal studies, where necessary	If confirmatory testing in vivo is deemed necessary for the assessment of genotoxicity, how can I design an animal study for optimal information gain?
4) Waiving or triggering of experimental studies based on bioavailability and distribution properties	How to avoid unnecessary testing, and how to determine when testing is needed, based on predictions of bioavailability and distribution (e.g. distribution to germ cells in genotoxicity assessment, distribution to the CNS in neurotoxicity assessment)?
5) Substantiating read-across arguments	I have a 90-day NOEL for chemical X, and need to assess the subchronic effects of an analogue, chemical Y. Can I read-across with confidence, based on both structural and toxicokinetic similarity?
6) Extrapolation from high to low doses taking into account non-monotonicity	How to characterise the low-dose effects of endocrine-active substances?
7) Modelling the effects of chemical mixtures, taking into account TK and TD interactions	<p>How can I account for synergistic/antagonistic effects due to cross-talk in regulatory pathways, or effects on cytochrome P450 induction/inhibition?</p> <p>How can I account for synergistic/antagonistic effects due to interactions on absorption and distribution properties?</p>
8) Use of TK information to adjust default safety factors in risk assessment	When can I reduce the default assessment factor, and when is a higher factor warranted?
9) Assessment of human bioaccumulation potential as a property of high concern	Ranking of environmental pollutants based on their expected bioaccumulation in humans

Table 12.2. Measured parameters and most commonly used in vitro systems for ADME[†]

Process	Measured Endpoint	Measured Parameter (unit)	In vitro method
Absorption	Permeation of external barriers	<u>Oral</u> P_{app} (cm·h ⁻¹)	Caco-2, PAMPA
		<u>Dermal</u> $P_{s,app}$ (cm·h ⁻¹)	skin-PAMPA, OECD TG 428
		<u>Inhalatory</u> P_{app} (cm·h ⁻¹) $K_{b,a}$ (dimensionless)	Human airway epithelium models, alveolar barrier models, cell lines (e.g. Calu-3) Head space model, head space SPME
Distribution	Permeation of internal barriers	$K_{t,b}$ (dimensionless)	Ultracentrifugation, ultrafiltration, rapid equilibrium dialysis (RED), solid phase microextraction (SPME)
		P_{app} (cm·h ⁻¹)	BBB model, cell line barrier model, co-cultures barrier model, organ culture barrier model, placenta barrier
		f_u (dimensionless)	Ultracentrifugation, rapid equilibrium dialysis (RED), solid phase microextraction (SPME)
Metabolism	Intrinsic or hepatic clearance Identification of metabolites	K_m (mmol·mL ⁻¹)	Liver: microsomal incubation, hepatocyte suspension incubation, plated hepatocyte incubation, HepaRG, subcellular incubations (S9, microsomal protein, cytosol)
		V_{max} (mmol·h ⁻¹)	Intestine: Microsomes, cytosol
		CL_{int} (mL·h ⁻¹)	Lung: Microsomes, cytosol, enterocytes, whole organ culture
Excretion	Irreversible removal (clearance) from circulating blood	f_u (dimensionless)	Ultracentrifugation, rapid equilibrium dialysis (RED), solid phase microextraction (SPME)
		$K_{b,a}$ (dimensionless)	Head space model
		GFR (mL/min)	Human physiology

Footnotes. P_{app} : apparent barrier permeability or penetration coefficient/constant, $K_{b,a}$: inhalation blood:air *distribution coefficient* (volatiles), $K_{t,b}$: tissue:blood partition coefficient, f_u : unbound or free fraction of the compound, K_m : Michaelis Menten constant, V_{max} : maximum metabolic rate, CL_{int} : intrinsic clearance **GFR**: glomerular filtration rate, [†] adapted from (Bessemers et al.)

Chapter 12. Table of References with Notes

Regulatory test guidelines and guidance documents	
European Medicines Agency (EMA). Concept paper on the need for revision of the position on the replacement of animal studies by in vitro models. Reference number: EMA/CHMP/SWP/169839/2011, available from: http://www.ema.europa.eu	The proposed Guideline will replace the Position on Replacement of Animal Studies by in vitro Methods (CPMP/SWP/728/95).
European Medicines Agency (EMA). Replacement of animal studies by in vitro models. Reference number: EMA/CPMP/SWP/728/95, available from http://www.ema.europa.eu	Position adapted by the Committee for Proprietary Medicinal Products (CPMP). The paper addresses the feasibility of replacing in vivo animal studies by in vitro investigations in the preclinical development of medical products.
European Medicines Agency (EMA). ICH: S 3 A: Toxicokinetics: A guidance for assessing systemic exposure in toxicology studies - Step 5. Reference number: CPMP/ICH/384/95, available from http://www.ema.europa.eu	<p>This Note for Guidance concerns toxicokinetics only with respect to the development of pharmaceutical products intended for use in human subjects. In this context, toxicokinetics is defined as the generation of pharmacokinetic data, either as an integral component in the conduct of non-clinical toxicity studies or in specially designed supportive studies, in order to assess systemic exposure. These data may be used in the interpretation of toxicology findings and their relevance to clinical safety issues.</p> <p>The Note for Guidance has been developed in order to provide an understanding of the meaning and application of toxicokinetics and to provide guidance on developing test strategies in toxicokinetics. The guidance highlights the need to integrate pharmacokinetics into toxicity testing, which should aid in the interpretation of the toxicology findings and promote rational study design development.</p>
European Food Safety Authority (EFSA). Guidance on Dermal Absorption. EFSA Journal 2012;10(4):2665 [30 pp.]. doi:10.2903/j.efsa.2012.2665. Available from: http://www.efsa.europa.eu	<p>Significant aspects of the new EFSA guidance include:</p> <ul style="list-style-type: none"> • Revised (lower) dermal absorption default values of 75% or 25%; • Defined criteria for many aspects of dermal absorption, intended to improve consistency of interpretation e.g. for extrapolating data between products and on approaches to the use of tape stripping data; • Support for dermal absorption values based on <i>in vitro</i> data alone.” • The agreement at the Standing Committee (May 2012) was that the new guidance will apply to all applications received after 30th November 2012.
European Food Safety Authority (EFSA). Guidance for submission for	TK studies are a prerequisite and should be conducted using the international

<p>food additive evaluations. EFSA Journal 2012; 10(7):2760 [60 pp.]. doi:10.2903/j.efsa.2012.2760. Available from: http://www.efsa.europa.eu</p>	<p>agreed guidelines</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2010. Guideline for Testing of Chemicals, 417, Toxicokinetics. Available from: http://www.oecd.org</p>	<p>This Test Guideline describes in vivo studies that provide information on mass balance, absorption, bioavailability, tissue distribution, metabolism, excretion, and basic toxicokinetic parameters [e.g. AUC], as well as supplemental approaches that may provide useful information on toxicokinetics. Information from toxicokinetic studies helps to relate concentration or dose to the observed toxicity and to understand its mechanism of toxicity. The test substance ("unlabelled" or "radiolabelled" forms) is normally administered by an oral route, but other routes of administration may be applicable. Single dose administration of the substance (preferably a minimum of two dose levels) may be adequate, but repeated dose may be needed in some circumstances. Toxicokinetic studies should preferably be carried out in the same species as that used in other toxicological studies performed with the substance (normally the rat, a minimum of 4 animals of one sex for each dose). Initial estimation of absorption can be achieved by mass balance determination, but further investigations such as intravenous (IV) administration and biliary excretion studies might be necessary. Bioavailability can be determined from plasma/blood kinetics of oral and IV groups. The percent of the total dose in tissues should at a minimum be measured at the termination of experiment, but additional time points may also be needed. Metabolites present at 5 % or greater of the administered dose should be identified. The rate and extent of excretion of the administered dose should be determined by measuring the percent recovered dose from urine, faeces and expired air.</p>
<p>Regulatory test guidelines and guidance documents</p>	
<p>Organisation for Economic Cooperation and Development (OECD). 2010. Guideline for Testing of Chemicals, 428, Skin absorption In vitro method. Available from: http://www.oecd.org</p>	<p>This Test method has been designed to provide information on absorption of a test substance, (ideally radiolabelled), applied to the surface of a skin sample separating the two chambers (a donor chamber and a receptor chamber) of a diffusion cell. Static and flow-through diffusion cells are both acceptable.</p> <p>Skin from human or animal sources can be used. Although viable skin is preferred, non-viable skin can also be used. The skin has been shown to have the capability to metabolise some chemicals during percutaneous absorption. In this case, metabolites of the test chemical may be analysed by appropriate methods.</p>

	<p>Normally more than one concentration of the test substance is used in typical formulations, spanning the realistic range of potential human exposures. The application should mimic human exposure, normally 1-5 mg/cm² of skin for a solid and up to 10 µl/cm² for liquids. The temperature must be constant because it affects the passive diffusion of chemicals. The absorption of a test substance during a given time period (normally 24h) is measured by analysis of the receptor fluid and the distribution of the test substance chemical in the test system and the absorption profile with time should be presented.</p>
<p>Organisation for Economic Cooperation and Development (OECD). 2010. Guideline for Testing of Chemicals, 427, Skin absorption In vivo method. Available from: http://www.oecd.org</p>	<p>The in vivo percutaneous absorption study set out in this Test Guideline provides the linkage necessary to extrapolate from oral studies when making safety assessments following dermal exposure. The in vivo method, described in this guideline, allows the determination of the penetration of the test substance through the skin into the systemic compartment.</p> <p>The test substance, preferably radiolabelled, is applied, for a fixed period of time, to the clipped skin of animals at one or more appropriate dose levels in the form of a representative in-use preparation. The rat is the most commonly used species. At least four animals of one sex should be used for each test preparation and each scheduled termination time. A known amount of the test preparation is evenly applied to the site. This amount should normally mimic potential human exposure, typically 1-5 mg/cm² for a solid or up to 10 µl/cm² for liquids. A relevant exposure period (typically 6 or 24 hours) should be used, based on the expected human exposure duration. The animals should be observed for signs of toxicity/abnormal reactions at intervals for the entire duration of the study. This study includes: daily measurements (excreta), regular detailed observations, as well as sacrifice at the scheduled time and blood collected for analysis.</p>
<p>World Health Organization (WHO). 2010. Characterization and application of physiologically based pharmacokinetic models in risk assessment. Available from: http://www.who.int/ipcs/methods/harmonization/areas/pbpb_models.pdf</p>	<p>A guidance document on "Principles of Characterizing and Applying PBPK Models in Risk Assessment".</p>

Chapter 12. Review papers from 2000

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Sung, J. H., Esch M. B., Prot J. M., Long C. J., Smith A., Hickman J. J., Shuler M. L., 2013. Microfabricated mammalian organ systems and their integration into models of whole animals and humans. Lab on a Chip - Miniaturisation for Chemistry and Biology. 13(7), 1201-1212	X	X	X	X	X	
Smith, D. A., 2013. Evolution of ADME science: Where else can modeling and simulation contribute? Molecular Pharmaceutics. 10(4), 1162-1170	X	X	X	X		X
Lanevskij, K., Japertas P., Didziapetris R., 2013. Improving the prediction of drug disposition in the brain. Expert Opinion on Drug Metabolism and Toxicology. 9(4), 473-486		X				X
Sarmiento, B., Andrade F., Da Silva S. B., Rodrigues F., Das Neves J., Ferreira D., 2012. Cell-based in vitro models for predicting drug permeability. Expert Opinion on Drug Metabolism and Toxicology. 8(5), 607-621	X				X	
Patel, B., Gauvin R., Absar S., Gupta V., Gupta N., Nahar K., Khademhosseini A., Ahsan F., 2012. Computational and bioengineered lungs as alternatives to whole animal, isolated organ, and cell-based lung models. American Journal of Physiology - Lung Cellular and Molecular Physiology. 303(9), L733-L747	X				X	X
Chen, Y., Jin J. Y., Mukadam S., Malhi V., Kenny J. R., 2012. Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: Strategy and approach during the drug discovery phase with four case studies. Biopharmaceutics and Drug Disposition. 33(2), 85-98	X	X		X		X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Buckley, S. T., Fischer S. M., Fricker G., Brandl M., 2012. In vitro models to evaluate the permeability of poorly soluble drug entities: Challenges and perspectives. <i>European Journal of Pharmaceutical Sciences</i> . 45(3), 235-250	X				X	
Braga, R. C., Andrade C. H., 2012. QSAR and QM/MM approaches applied to drug metabolism prediction. <i>Mini-Reviews in Medicinal Chemistry</i> . 12(6), 573-582			X			X
Bell, L. C., Wang J., 2012. Probe ADME and Test Hypotheses: A PATH beyond clearance in vitro in vivo correlations in early drug discovery. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 8(9), 1131-1155			X		X	X
Bansal, P., Ajay D., 2012. Laboratory dialysis-past, present and future. <i>Recent Patents on Biotechnology</i> . 6(1), 32-44		X			X	
Andersson, T. B., Kanebratt K. P., Kenna J. G., 2012. The HepaRG cell line: A unique in vitro tool for understanding drug metabolism and toxicology in human. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 8(7), 909-920			X		X	
Wu, Z., Zheng F., Ding L., 2011. Development of in vitro evaluation methods to access membrane permeability of drugs. <i>Journal of China Pharmaceutical University</i> . 42(1), 16-21	X				X	
Wu, Y. F., Liu H., Ni J. M., 2011. Advances in parallel artificial membrane permeability assay and its applications. <i>Yaoxue Xuebao</i> . 46(8), 890-895	X				X	
Van Midwoud, P. M., Verpoorte E., Groothuis G. M. M., 2011. Microfluidic devices for in vitro studies on liver drug metabolism and toxicity. <i>Integrative Biology</i> . 3(5), 509-521			X		X	
Tóth, A., Veszélka S., Nakagawa S., Niwa M., Deli M. A., 2011. Patented in vitro blood-brain barrier models in CNS drug discovery. <i>Recent Patents on CNS Drug Discovery</i> . 6(2), 107-118		X			X	

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Tarcsay, A.,Keserü G. M., 2011. In silico site of metabolism prediction of cytochrome P450-mediated biotransformations. Expert Opinion on Drug Metabolism and Toxicology. 7(3), 299-312			X			X
Smetanová, L., Stětinová V., Svoboda Z.,Kvetina J., 2011. Caco-2 cells, biopharmaceutics classification system (BCS) and biowaiver. Acta medica (Hradec Králové) / Universitas Carolina, Facultas Medica Hradec Králové. 54(1), 3-8	X				X	
Schroeder, K., Bremm K. D., Alépée N., Bessems J. G. M., Blaauboer B., Boehn S. N., Burek C., Coecke S., Gombau L., Hewitt N. J., Heylings J., Huwyler J., Jaeger M., Jagelavicius M., Jarrett N., Ketelslegers H., Kocina I., Koester J., Kreysa J., Note R., Poth A., Radtke M., Rogiers V., Scheel J., Schulz T., Steinkellner H., Toeroek M., Whelan M., Winkler P.,Diembeck W., 2011. Report from the EPAA workshop: In vitro ADME in safety testing used by EPAA industry sectors. Toxicology in Vitro. 25(3), 589-604	X	X	X	X	X	
Reichl, S., Kölln C., Hahne M.,Verstraelen J., 2011. In vitro cell culture models to study the corneal drug absorption. Expert Opinion on Drug Metabolism and Toxicology. 7(5), 559-578	X				X	
Peters, F. T.,Meyer M. R., 2011. In vitro approaches to studying the metabolism of new psychoactive compounds. Drug Testing and Analysis. 3(7-8), 483-495			X		X	
Mudra, D. R., Desino K. E.,Desai P. V., 2011. In silico, in vitro and in situ models to assess interplay between CYP3A and P-gp. Current Drug Metabolism. 12(8), 750-773			X		X	X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Mandenius, C. F., Andersson T. B., Alves P. M., Batzl-Hartmann C., Björquist P., Carrondo M. J. T., Chesne C., Coecke S., Edsbagge J., Fredriksson J. M., Gerlach J. C., Heinzle E., Ingelman-Sundberg M., Johansson I., Küppers-Munther B., Müller-Vieira U., Noor F., Zeilinger K., 2011. Toward preclinical predictive drug testing for metabolism and hepatotoxicity by using in vitro models derived from human embryonic stem cells and human cell lines - A report on the vitroculture EU-project. ATLA Alternatives to Laboratory Animals. 39(2), 147-171			X		X	
Lušin, T. T., Trontelj J., 2011. In vitro models for investigation of drug metabolism in the liver. In vitro modeli za preučevanje metabolizma učinkovin v jetrih. 62(4), 171-174			X		X	
Li, H. F., Zou J., Bai R. Y., Xing Y. M., Nie J. M., Diao Y., 2011. M cell in vitro model and its application in oral delivery of macromolecular drugs. Yaoxue Xuebao. 46(12), 1429-1435	X				X	
Hecht, D., 2011. Applications of machine learning and computational intelligence to drug discovery and development. Drug Development Research. 72(1), 53-65	X	X	X	X		X
Wang, J., Skolnik S., 2010. Mitigating permeability-mediated risks in drug discovery. Expert Opinion on Drug Metabolism and Toxicology. 6(2), 171-187	X				X	X
Vuignier, K., Schappler J., Veuthey J. L., Carrupt P. A., Martel S., 2010. Drug-protein binding: A critical review of analytical tools. Analytical and Bioanalytical Chemistry. 398(1), 53-66		X		X	X	
Sung, J. H., Esch M. B., Shuler M. L., 2010. Integration of in silico and in vitro platforms for pharmacokinetic pharmacodynamic modeling. Expert Opinion on Drug Metabolism and Toxicology. 6(9), 1063-1081	X	X	X		X	X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Lademann, J., Richter H., Meinke M., Sterry W., Patzelt A., 2010. Which skin model is the most appropriate for the investigation of topically applied substances into the hair follicles? <i>Skin Pharmacology and Physiology</i> . 23(1), 47-52	X				X	
Kharkar, P. S., 2010. Two-dimensional (2D) in silico models for Absorption, Distribution, Metabolism, Excretion and Toxicity (ADME/T) in drug discovery. <i>Current Topics in Medicinal Chemistry</i> . 10(1), 116-126	X	X	X	X		X
Habka, D., Péry A. R. R., Legallais C., Brochot C., 2010. Predicting in vivo toxicokinetics of chemicals from in vitro data and QSAR models. Prédiction de la toxicocinétique in vivo de substances chimiques à partir de données in vitro et de modèles QSAR. 9(6), 489-501	X	X	X	X		X
Burton, P. S., Goodwin J. T., 2010. Solubility and permeability measurement and applications in drug discovery. <i>Combinatorial Chemistry and High Throughput Screening</i> . 13(2), 101-111	X					X
Asha, S., Vidyavathi M., 2010. Role of human liver microsomes in in vitro metabolism of drugs-A review. <i>Applied Biochemistry and Biotechnology</i> . 160(6), 1699-1722			X		X	
Álvarez-Lueje, A., Dragnic S. B., 2010. Stability and drug metabolism assessed by electrochemical methods. <i>Combinatorial Chemistry and High Throughput Screening</i> . 13(8), 713-727			X		X	
Wang, J., Hou T., 2009. Chapter 5 Recent Advances on in silico ADME Modeling. 5(101-127)	X	X	X	X		X
Vastag, M., Keseru G. M., 2009. Current in vitro and in silico models of blood-brain barrier penetration: A practical view. <i>Current Opinion in Drug Discovery and Development</i> . 12(1), 115-124		X			X	X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Valerio Jr, L. G., 2009. In silico toxicology for the pharmaceutical sciences. Toxicology and Applied Pharmacology. 241(3), 356-370	X	X	X	X		X
Missner, A.,Pohl P., 2009. 110 years of the Meyer-Overton rule: Predicting membrane permeability of gases and other small compounds. ChemPhysChem. 10(9-10), 1405-1414	X					X
Korjamo, T., Heikkinen A. T.,Mönkkönen J., 2009. Analysis of unstirred water layer in in vitro permeability experiments. Journal of Pharmaceutical Sciences. 98(12), 4469-4479	X				X	
Klon, A. E., 2009. Computational models for central nervous system penetration. Current Computer-Aided Drug Design. 5(2), 71-89		X				X
Jacob, A., Pratuangdejkul J., Buffet S., Launay J. M.,Manivet P., 2009. In silico platform for xenobiotics ADME-T pharmacological properties modeling and prediction. Part II: the body in a Hilbertian space. Drug Discovery Today. 14(7-8), 406-412	X	X	X	X		X
Huynh, L., Masereeuw R., Friedberg T., Ingelman-Sundberg M.,Manivet P., 2009. In silico platform for xenobiotics ADME-T pharmacological properties modeling and prediction. Part I: beyond the reduction of animal model use. Drug Discovery Today. 14(7-8), 401-405	X	X	X	X		X
Huong, S. P., Bun H., Fourneron J. D., Reynier J. P.,Andrieu V., 2009. Use of various models for in vitro percutaneous absorption studies of ultraviolet filters. Skin Research and Technology. 15(3), 253-261	X				X	
Hage, D. S., Jackson A., Sobansky M. R., Schiel J. E., Yoo M. J.,Joseph K. S., 2009. Characterization of drug-protein interactions in blood using high-performance affinity chromatography. Journal of Separation Science. 32(5-6), 835-853		X		X	X	

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Franklin, R. B., 2009. In Silico studies in ADME/Tox: Caveat emptor. Current Computer-Aided Drug Design. 5(2), 128-138	X	X	X	X		X
D'Souza, M. J., Koyoshi F., Everett L., 2009. Structure Activity Relationships (SARs) using a structurally diverse drug database: Validating success of predictor tools. Pharmaceutical Reviews. 7(5),	X	X	X	X		X
Czodrowski, P., Kriegl J. M., Scheuerer S., Fox T., 2009. Computational approaches to predict drug metabolism. Expert Opinion on Drug Metabolism and Toxicology. 5(1), 15-27			X			X
Barbero, A. M., Frasc H. F., 2009. Pig and guinea pig skin as surrogates for human in vitro penetration studies: A quantitative review. Toxicology in Vitro. 23(1), 1-13	X				X	
Wang, L., Chen H., 2008. A screening model of drug absorption, distribution, metabolism and excretion in vitro. Journal of Clinical Rehabilitative Tissue Engineering Research. 12(50), 9957-9961	X	X	X	X	X	
Varnek, A., Fourches D., Horvath D., Klimchuk O., Gaudin C., Vayer P., Solov'ev V., Hoonakker F., Tetko I. V., Marcou G., 2008. ISIDA - Platform for virtual screening based on fragment and pharmacophoric descriptors. Current Computer-Aided Drug Design. 4(3), 191-198	X	X	X	X		X
Sun, J., Wu X., Lu R., Liu J., Wang Y., He Z., 2008. Profiling drug membrane permeability and activity via biopartitioning chromatography. Current Drug Metabolism. 9(2), 152-166	X				X	
Pichandy, M., Kannaiyan S., Mandhava V., Sindhu B., Lakshmi P. K., Umadevi, Vijayanthi, Mithra A., 2008. Various in vitro models to assess drug permeability across the blood brain barrier. Biosciences Biotechnology Research Asia. 5(1), 433-437		X			X	

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Martins-Green, M., Petreaca M., Yao M., 2008. Chapter 8 An Assay System for In Vitro Detection of Permeability in Human "Endothelium". 443(137-153)	X				X	
Lipscomb, J. C., Poet T. S., 2008. In vitro measurements of metabolism for application in pharmacokinetic modeling. Pharmacology and Therapeutics. 118(1), 82-103			X		X	
Koukoulitsa, C., 2008. In silico approaches for the discovery of active leads from nature. Pharmakeftiki. 21(2), 57-62	X	X	X	X		X
Kirchmair, J., Distinto S., Schuster D., Spitzer G., Langer T., Wolber G., 2008. Enhancing drug discovery through in silico screening: Strategies to increase true positives retrieval rates. Current Medicinal Chemistry. 15(20), 2040-2053	X	X	X	X		X
Jin, H., Di L., 2008. Permeability - In vitro assays for assessing drug transporter activity. Current Drug Metabolism. 9(9), 911-920	X				X	
Jakasa, I., Kezic S., 2008. Evaluation of in-vivo animal and in-vitro models for prediction of dermal absorption in man. Human and Experimental Toxicology. 27(4), 281-288	X				X	
Hou, T., Wang J., 2008. Structure-ADME relationship: Still a long way to go? Expert Opinion on Drug Metabolism and Toxicology. 4(6), 759-770	X	X	X	X		X
Hop, C. E. C. A., Cole M. J., Davidson R. E., Duignan D. B., Federico J., Janiszewski J. S., Jenkins K., Krueger S., Lebowitz R., Liston T. E., Mitchell W., Snyder M., Steyn S. J., Soglia J. R., Taylor C., Troutman M. D., Umland J., West M., Whalen K. M., Zelesky V., Zhao S. X., 2008. High throughput ADME screening: Practical considerations, impact on the portfolio and enabler of in silico ADME models. Current Drug Metabolism. 9(9), 847-853	X	X	X	X	X	

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Faller, B., 2008. Artificial membrane assays to assess permeability. <i>Current Drug Metabolism</i> . 9(9), 886-892	X				X	
Donato, M. T., Lahoz A., Castell J. V., Gómez-Lechón M. J., 2008. Cell lines: A tool for in vitro drug metabolism studies. <i>Current Drug Metabolism</i> . 9(1), 1-11			X		X	
De La Nuez, A., Rodríguez R., 2008. Current methodology for the assessment of ADME-Tox properties on drug candidate molecules. <i>Biotechnologia Aplicada</i> . 25(2), 97-110	X	X	X	X		X
Chohan, K. K., Paine S. W., Waters N. J., 2008. Advancements in predictive in silico models for ADME. <i>Current Chemical Biology</i> . 2(3), 215-228	X	X	X	X		X
Carlson, T. J., Fisher M. B., 2008. Recent advances in high throughput screening for ADME properties. <i>Combinatorial Chemistry and High Throughput Screening</i> . 11(3), 258-264	X	X	X	X	X	
Banker, M. J., Clark T. H., 2008. Plasma/serum protein binding determinations. <i>Current Drug Metabolism</i> . 9(9), 854-859		X		X	X	
Abbott, N. J., Dolman D. E. M., Patabendige A. K., 2008. Assays to predict drug permeation across the blood-brain barrier, and distribution to brain. <i>Current Drug Metabolism</i> . 9(9), 901-910		X			X	
Xia, C. Q., Milton M. N., Gan L. S., 2007. Evaluation of drug-transporter interactions using in vitro and in vivo models. <i>Current Drug Metabolism</i> . 8(4), 341-363		X			X	
Wan, H., Bergström F., 2007. High throughput screening of drug-protein binding in drug discovery. <i>Journal of Liquid Chromatography and Related Technologies</i> . 30(5-7), 681-700		X		X	X	
van de Kerkhof, E. G., de Graaf I. A. M., Groothuis G. M. M., 2007. In vitro methods to study intestinal drug metabolism. <i>Current Drug Metabolism</i> . 8(7), 658-675			X		X	

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Sumantran, V. N., 2007. Experimental approaches for studying uptake and action of herbal medicines. <i>Phytotherapy Research</i> . 21(3), 210-214	X				X	
Nigsch, F., Klaffke W., Miret S., 2007. In vitro models for processes involved in intestinal absorption. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 3(4), 545-556	X				X	
Jolivet, L. J., Ekins S., 2007. Methods for Predicting Human Drug Metabolism. 43(131-176)			X		X	X
Jiang, H., An P. L., Jiang Y., 2007. Drug and protein interactions: Progress of research methods. <i>Academic Journal of Second Military Medical University</i> . 28(6), 662-666		X		X	X	
Gómez-Lechón, M. J., Donato M. T., 2007. In vitro investigation of drug metabolism and toxicity in man. <i>Anales de la Real Academia Nacional de Farmacia</i> . 73(1), 5-26			X		X	
Galini-Luciani, D., Nguyen L., Yazdani M., 2007. Is PAMPA a useful tool for discovery? <i>Journal of Pharmaceutical Sciences</i> . 96(11), 2886-2892	X				X	
Fortuna, A., Alves G., Falcão A., 2007. The importance of permeability screening in drug discovery process: PAMPA, Caco-2 and rat everted gut assays. <i>Current Topics in Pharmacology</i> . 11(2), 63-86	X				X	
De Souza, J., Freitas Z. M. F., Storpirtis S., 2007. In vitro models for the determination of drug absorption and a prediction of dissolution/absorption relationships. Modelos in vitro para determinação da absorção de fármacos e previsão da relação dissolução/absorção. 43(4), 515-527	X				X	
Yap, C. W., Xue Y., Li Z. R., Chen Y. Z., 2006. Application of support vector machines to in silico prediction of cytochrome P450 enzyme substrates and inhibitors. <i>Current Topics in Medicinal Chemistry</i> . 6(15), 1593-1607			X			X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Sakagami, M., 2006. In vivo, in vitro and ex vivo models to assess pulmonary absorption and disposition of inhaled therapeutics for systemic delivery. <i>Advanced Drug Delivery Reviews</i> . 58(9-10), 1030-1060	X				X	
Madden, J. C., Cronin M. T. D., 2006. Structure-based methods for the prediction of drug metabolism. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 2(4), 545-557			X		X	
Gombar, V. K., Alberts J. J., Cassidy K. C., Mattioni B. E., Mohutsky M. A., 2006. In silico metabolism studies in drug discovery: Prediction of metabolic stability. <i>Current Computer-Aided Drug Design</i> . 2(2), 177-188			X			X
Bidault, Y., 2006. A flexible approach for optimising in silico ADME/Tox characterisation of lead candidates. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 2(1), 157-168	X	X	X	X		X
Balimane, P. V., Han Y. H., Chong S., 2006. Current industrial practices of assessing permeability and P-glycoprotein interaction. <i>AAPS Journal</i> . 8(1),	X				X	
Willmann, S., Lippert J., Schmitt W., 2005. From physicochemistry to absorption and distribution: Predictive mechanistic modelling and computational tools. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 1(1), 159-168	X	X				X
Williams, J. A., Bauman J., Cai H., Conlon K., Hansel S., Hurst S., Sadagopan N., Tugnait M., Zhang L., Sahi J., 2005. In vitro ADME phenotyping in drug discovery: Current challenges and future solutions. <i>Current Opinion in Drug Discovery and Development</i> . 8(1), 78-88	X	X	X	X		
Subramanian, K., 2005. TruPK - Human pharmacokinetic models for quantitative ADME prediction. <i>Expert Opinion on Drug Metabolism and Toxicology</i> . 1(3), 555-564	X	X	X	X		

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	IN VITRO	IN SILICO
Steimer, A., Haltner E.,Lehr C. M., 2005. Cell culture models of the respiratory tract relevant to pulmonary drug delivery. <i>Journal of Aerosol Medicine: Deposition, Clearance, and Effects in the Lung.</i> 18(2), 137-182	X				X	
Sivaraman, A., Leach J. K., Townsend S., Iida T., Hogan B. J., Stolz D. B., Fry R., Samson L. D., Tannenbaum S. R.,Griffith L. G., 2005. A microscale in vitro physiological model of the liver: Predictive screens for drug metabolism and enzyme induction. <i>Current Drug Metabolism.</i> 6(6), 569-591			X		X	
Hashida, M., 2005. In silico prediction of pharmacokinetic properties. <i>Yakugaku Zasshi.</i> 125(11), 853-861	X	X	X	X		X
Castell, J. V., Donato M. T.,Gómez-Lechón M. J., 2005. Metabolism and bioactivation of toxicants in the lung. The in vitro cellular approach. <i>Experimental and Toxicologic Pathology.</i> 57(SUPPL. 1), 189-204			X		X	
Balimane, P. V.,Chong S., 2005. Cell culture-based models for intestinal permeability: A critique. <i>Drug Discovery Today.</i> 10(5), 335-343	X				X	
Balakin, K. V.,Ivanenkov Y. A., 2005. Computational methods for prediction of metabolism of physiologically active substances. <i>Biomeditsinskaya Khimiya.</i> 51(4), 384-412			X			X
Zhou, D., Li L.,Li F., 2004. Analytical methodology for drug-protein binding studies. <i>Se pu = Chinese journal of chromatography / Zhongguo hua xue hui.</i> 22(2), 116-120		X		X	X	
Yamashita, F.,Hashida M., 2004. In silico approaches for predicting ADME properties of drugs. <i>Drug metabolism and pharmacokinetics.</i> 19(5), 327-338	X	X	X	X		X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
van der Bijl, P., van Eyk A. D., 2004. Human vaginal mucosa as a model of buccal mucosa for in vitro permeability studies: an overview. <i>Current drug delivery</i> . 1(2), 129-135		X			X	
Rooyackers, O., Thorell A., Nygren J., Ljungqvist O., 2004. Microdialysis methods for measuring human metabolism. <i>Current Opinion in Clinical Nutrition and Metabolic Care</i> . 7(5), 515-521			X		X	
Mälkiä, A., Murtomäki L., Urtti A., Kontturi K., 2004. Drug permeation in biomembranes: In vitro and in silico prediction and influence of physicochemical properties. <i>European Journal of Pharmaceutical Sciences</i> . 23(1), 13-47	X	X			X	X
Langer, O., Müller M., 2004. Methods to assess tissue-specific distribution and metabolism of drugs. <i>Current Drug Metabolism</i> . 5(6), 463-481		X	X		X	
Kansy, M., Avdeef A., Fischer H., 2004. Advances in screening for membrane permeability: High-resolution PAMPA for medicinal chemists. <i>Drug Discovery Today: Technologies</i> . 1(4), 349-355	X				X	
Harris, D. L., 2004. In silico predictive metabolism: A structural/electronic filter method. <i>Current Opinion in Drug Discovery and Development</i> . 7(1), 43-48			X			X
Dickins, M., Van De Waterbeemd H., 2004. Simulation models for drug disposition and drug interactions. <i>Drug Discovery Today: BIOSILICO</i> . 2(1), 38-45	X	X	X	X		X
Beresford, A. P., Segall M., Tarbit M. H., 2004. In silico prediction of ADME properties: Are we making progress? <i>Current Opinion in Drug Discovery and Development</i> . 7(1), 36-42	X	X	X	X		X
Zhang, Y., Bachmeier C., Miller D. W., 2003. In vitro and in vivo models for assessing drug efflux transporter activity. <i>Advanced Drug Delivery Reviews</i> . 55(1), 31-51	X	X		X	X	X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Yu, H.,Adedoyin A., 2003. ADME-Tox in drug discovery: Integration of experimental and computational technologies. <i>Drug Discovery Today</i> . 8(18), 852-861	X	X	X	X	X	X
van de Waterbeemd, H.,Gifford E., 2003. ADMET in silico modelling: Towards prediction paradise? <i>Nature Reviews Drug Discovery</i> . 2(3), 192-204	X	X	X	X		X
Tsai, T. H., 2003. Assaying protein unbound drugs using microdialysis techniques. <i>Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences</i> . 797(1-2), 161-173		X		X	X	
Lombardo, F., Gifford E.,Shalaeva M. Y., 2003. In silico ADME prediction: data, models, facts and myths. <i>Mini reviews in medicinal chemistry</i> . 3(8), 861-875	X	X	X	X		X
Heringa, M. B.,Hermens J. L. M., 2003. Measurement of free concentrations using negligible depletion-solid phase microextraction (nd-SPME). <i>TrAC - Trends in Analytical Chemistry</i> . 22(9), 575-587		X		X	X	
Handschin, C., Podvinec M.,Meyer U. A., 2003. In Silico Approaches, and In Vitro and In Vivo Experiments to Predict Induction of Drug Metabolism. <i>Drug News and Perspectives</i> . 16(7), 423-434			X		X	X
Gombar, V. K., Silver I. S.,Zhao Z., 2003. Role of ADME characteristics in drug discovery and their in silico evaluation: in silico screening of chemicals for their metabolic stability. <i>Current topics in medicinal chemistry</i> . 3(11), 1205-1225	X	X	X	X		X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Gebhardt, R., Hengstler J. G., Müller D., Glöckner R., Buenning P., Laube B., Schmelzer E., Ullrich M., Utesch D., Hewitt N., Ringel M., Hilz B. R., Bader A., Langsch A., Koose T., Burger H. J., Maas J., Oesch F., 2003. New hepatocyte in vitro systems for drug metabolism: Metabolic capacity and recommendations for application in basic research and drug development, standard operation procedures. <i>Drug Metabolism Reviews</i> . 35(2-3), 145-213			X		X	
Fleishaker, J. C., 2003. Models and methods for predicting drug transfer into human milk. <i>Advanced Drug Delivery Reviews</i> . 55(5), 643-652				X	X	X
Colmenarejo, G., 2003. In silico prediction of drug-binding strengths to human serum albumin. <i>Medicinal Research Reviews</i> . 23(3), 275-301		X		X		X
Van de Waterbeemd, H., 2002. High-throughput and in silico techniques in drug metabolism and pharmacokinetics. <i>Current Opinion in Drug Discovery and Development</i> . 5(1), 33-43			X		X	X
Lundquist, S., Renftel M., 2002. The use of in vitro cell culture models for mechanistic studies and as permeability screens for the blood-brain barrier in the pharmaceutical industry - Background and current status in the drug discovery process. <i>Vascular Pharmacology</i> . 38(6), 355-364		X			X	
Langowski, J., Long A., 2002. Computer systems for the prediction of xenobiotic metabolism. <i>Advanced Drug Delivery Reviews</i> . 54(3), 407-415			X			X
Darvas, F., Keseru G., Papp A., Dormán G., Urge L., Krajcsi P., 2002. In Silico and Ex silico ADME approaches for drug discovery. <i>Current topics in medicinal chemistry</i> . 2(12), 1287-1304	X	X	X	X		X
Butina, D., Segall M. D., Frankcombe K., 2002. Predicting ADME properties in silico: Methods and models. <i>Drug Discovery Today</i> . 7(11), S83-S88	X	X	X	X		X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Boyer, S.,Zamora I., 2000. New methods in predictive metabolism. Molecular Diversity. 5(4), 277-287			X			X
Borchard, G., 2002. Calu-3 cells, a valid model for the airway epithelium? S.T.P. Pharma Sciences. 12(4), 205-211	X		X		X	
Balls, M., Bogni A., Bremer S., Casati S., Coecke S., Eskes C., Prieto P., Sabbioni E., Worth A., Zuang V., Barratt M., Blauboer B., Botham P., Combes R., Doehmer J., Fentem J., Liebsch M.,Spielmann H., 2002. Alternative (non-animal) methods for chemicals testing: Current status and future prospects. ATLA Alternatives to Laboratory Animals. 30(SUPPL. 1), i-vi+1-125	X	X	X	X	X	X
Agu, R. U., Jorissen M., Kinget R., Verbeke N.,Augustijns P., 2002. Alternatives to in vivo nasal toxicological screening for nasally-administered drugs. S.T.P. Pharma Sciences. 12(1), 13-22	X				X	
Venkatakrisnan, K., Von Moltke L. L.,Greenblatt D. J., 2001. Human drug metabolism and the cytochromes P450: Application and relevance of in vitro models. Journal of Clinical Pharmacology. 41(11), 1149-1179			X		X	
Podlogar, B. L.,Muegge I., 2001. "Holistic" in silico methods to estimate the systemic and CNS bioavailabilities of potential chemotherapeutic agents. Current topics in medicinal chemistry. 1(4), 257-275	X		X			X
Gumbleton, M.,Audus K. L., 2001. Progress and limitations in the use of in vitro cell cultures to serve as a permeability screen for the blood-brain barrier. Journal of Pharmaceutical Sciences. 90(11), 1681-1698		X			X	
Bohets, H., Annaert P., Mannens G., Van Beijsterveldt L., Anciaux K., Verboven P., Meuldermans W.,Lavrijsen K., 2001. Strategies for absorption screening in drug discovery and development. Current topics in medicinal chemistry. 1(5), 367-383	X				X	X

TITLE	ABSORPTION	DISTRIBUTION	METABOLISM	EXCRETION	<i>IN VITRO</i>	<i>IN SILICO</i>
Sartorelli, P., Andersen H. R., Angerer J., Corish J., Drexler H., Göen T., Griffin P., Hotchkiss S. A. M., Larese F., Montomoli L., Perkins J., Schmelz M., Van De Sandt J., Williams F., 2000. Percutaneous penetration studies for risk assessment. <i>Environmental Toxicology and Pharmacology</i> . 8(2), 133-152	X				X	
Hengstler, J. G., Utesch D., Steinberg P., Platt K. L., Diener B., Ringel M., Swales N., Fischer T., Biefang K., Gerl M., Böttger T., Oesch F., 2000. Cryopreserved primary hepatocytes as a constantly available in vitro model for the evaluation of human and animal drug metabolism and enzyme induction. <i>Drug Metabolism Reviews</i> . 32(1), 81-118			X		X	
Forbes, B., 2000. Human airway epithelial cell lines for in vitro drug transport and metabolism studies. <i>Pharmaceutical Science and Technology Today</i> . 3(1), 18-27	X		X		X	
Boyer, S., Zamora I., 2000. New methods in predictive metabolism. <i>Molecular Diversity</i> . 5(4), 277-287			X			X
Prieto, P., B. J. Blauboer, A. G. De Boer, M. Boveri, R. Cecchelli, C. Clemenson, S. Coecke, A. Forsby, H. J. Galla, P. Garberg, J. Greenwood, A. Price and H. Tähti 2004. Blood-Brain Barrier In Vitro Models and Their Application in Toxicology: The Report and Recommendations of ECVAM Workshop 49. <i>ATLA Alternatives to Laboratory Animals</i> 32(1): 37-50.		X			X	

13. Aquatic (fish) toxicity

Marlies Halder, Aude Kienzler & Andrew Worth

13.1 Introduction

Aquatic toxicity refers to the effects of a compound to organisms living in the water compartment and is usually determined by testing on organisms representing the three trophic levels, i.e. plants (algae), invertebrates (crustaceans as *Daphnia spp.*) and vertebrates (fish). This chapter focuses on standard and non-standard approaches for assessing fish toxicity (excluding bioaccumulation and endocrine disruptors). The two toxicological endpoints using fish are: a) acute fish toxicity (short-term exposure to determine the concentration that is lethal to 50% of the fish, LC50); and b) chronic fish toxicity (long-term exposure covering the life-cycle of the fish, to identify sublethal effects and determine the No Observed Effect Concentration [NOEC], the Lowest Observed Effect Concentration [LOEC], or effective concentration [EC_x]).

13.2 The traditional animal tests and their regulatory use

The OECD Test Guidelines (TGs) / EU test methods for aquatic toxicity involve the use of fish at various life stages. For short-term toxicity, juvenile or adult fish are used, whereas long-term toxicity testing also covers early life stages: fertilised egg – embryo – (hatch) – sac-fry or larva (sac = yolk) – juvenile – adult. The duration of these developmental steps depends on the species used and the water temperature. Table 13.1 gives an overview on the OECD TGs and Table 13.2 summarises the regulatory requirements in the EU.

13.2.1 Acute (short-term) fish toxicity

The *fish acute toxicity test* (OECD TG 203) (OECD, 1992) is carried out with juvenile or adult fish. It is a short-term exposure test (96 h) and determines the concentration that is lethal to 50% of the fish (LC50). Other relevant endpoints, consistent with OECD TG 203, can include the LC0 and LC100 (0% mortality and 100% mortality). Groups of seven to 10 fish are exposed via static, semi-static, or flow-through systems to at least five concentrations of the test substance, a water control and, if needed, a solvent control. The mortality is recorded on a daily basis for each concentration and used for the LC50 calculation. In addition, any sublethal effects observed should be reported. Analytical confirmation of the exposure concentrations is recommended. OECD TG 203 also allows to carry out a *limit test* at a single concentration (= 100 mg/L) with seven to 10 fish in the treatment group and seven to 10 fish in the control / solvent group. If no mortality occurs, LC50 is reported as >100mg/L. If mortality occurs, the full test needs to be conducted.

Rufli and Springer (2011) demonstrated by retrospective analysis of acute fish toxicity tests that the number of fish / concentration could be reduced to six without loss of statistical power. In the interest of animal welfare, Rufli (2012) recommend that moribund (a state close to death involving severe suffering) fish be removed from tests and humanely killed. However, before introducing this modification into OECD TG 203, further analysis is necessary, since introducing “moribund stage” may lower LC50 values. This is the basis of a project in the OECD Test Guidelines programme.

Depending on the regulatory framework and geographic region, the fish species to be used varies, or the tests have to be carried out on cold and/or warm water species. The following are the most commonly used OECD species: zebrafish (*Danio rerio*), fathead minnow

(*Pimephales promelas*), medaka (*Oryzias latipes*), bluegill sunfish (*Lepomis macrochirus*), and rainbow trout (*Oncorhynchus mykiss*). In addition, the common carp (*Cyprinus carpio*) and guppy (*Poecilia reticulata*) are recommended species.

Another short-term fish toxicity test is described in OECD TG 204 (OECD, 1984), the *fish prolonged toxicity test – 14 days study*, which is carried out with juvenile or adult fish. In addition to lethal effects, sublethal effects are recorded and used for NOEC determination. Since this test is rarely conducted and OECD TG 203 includes the provision to extend the exposure beyond 96 h if necessary, e.g. in case of slow onset of mortality, the OECD Fish Toxicity Testing Framework report (OECD, 2012a) recommended deletion of this guideline. This recommendation was followed up by the OECD in late 2012 and the deletion of the guideline will come into force in April 2014.

Requirements for acute fish toxicity testing vary depending on the type of regulation and the geographic region (OECD, 2012a). In Europe, acute fish toxicity data are required for (Table 13.2):

- a) industrial chemicals (>10 t/year; EC, 2006);
- b) biocides; use of the threshold approach (see below) is recommended (EU, 2012a);
- c) plant protection products; for the active substance, data from rainbow trout and one warm water species; and for formulations, if fish is the most sensitive species for the active substance (EC, 2009a); from 2014, the threshold approach (see below) is recommended and only data from rainbow trout are required (EU, 2013a; 2013b)
- d) veterinary pharmaceuticals (EMA, 2004), and
- e) others (e.g. feed; EC, 2008; EFSA, 2008).

Aquatic toxicity is not an endpoint in the cosmetics regulation, however, environmental concerns of cosmetics ingredients and products are considered through REACH (EC, 2009b).

Annex VIII of the REACH regulation (EC, 2006) states that a study for short-term aquatic toxicity does not need to be conducted if “*there are mitigating factors indicating that aquatic toxicity is unlikely to occur, for instance if the substance is highly insoluble in water or the substance is unlikely to cross biological membranes.*” Moreover, “*applicants are invited to consider aquatic long-term aquatic toxicity testing.*” Long-term toxicity testing should be considered at this tonnage level in two cases: a) “*if the chemical safety assessment according to Annex I indicates the need to investigate further the effects on aquatic organisms*” (i.e. if PEC/PNEC >1 and for chemicals with log Kow >3 [or BCF >100] and a PEC >1/100th of the water solubility) and b) “*if the substance is poorly water soluble.*”

13.2.2 Chronic (long-term) fish toxicity

Chronic / long-term fish toxicity tests cover several life stages of fish depending on the OECD TG chosen.

The most commonly used test to establish chronic fish toxicity is OECD TG 210, the *fish early life-stage (FELS) test*. The initial guideline of 1992 was reviewed and recently approved (OECD, 2013a). OECD TG 210 aims to determine lethal and sublethal effects of a chemical on early life stages of fish (embryos, larvae, juvenile fish). The endpoints are hatching success, abnormal appearance, abnormal behaviour, survival / mortality, and the weight and length of the fish at the end of the test. By comparison to the values of control fish, the LOEC, NOEC and/or ECx are determined for each endpoint. The test starts with fertilised eggs and continues until the control fish are freely feeding. In general, at least 5

concentrations with at least 80 eggs per concentration divided in four replicates are used as well as a control group and a solvent control group if needed. OECD TG 210 recommends four freshwater species (zebrafish, fathead minnow, medaka, rainbow trout) and a saltwater species (sheephead minnow). The revised OECD TG 210 allows the use of a limit test, or an extended limit test, with fewer than five concentrations (to be justified) (OECD, 2013a).

OECD TG 212, *Fish, Short-term Toxicity Tests on Embryo and Sac-fry Stages* (OECD, 1998), is less frequently carried out and covers embryo and sac-fry stages. Exposure starts with fertilised eggs and should be terminated before the yolk-sac is completely absorbed or before mortality due to starvation begins. It assesses effects of a chemical on hatching, abnormal appearance, abnormal behaviour, survival / mortality (at the embryo and sac-fry stages and overall), and the weight and length at the end of the test. The NOEC and LOEC are determined for the endpoints observed. The duration of the test varies depending on the species, e.g. for zebrafish 8-10 days are given in OECD TG 212. There is a major animal welfare concern associated with this guideline, which has also been referred to as the fish “starvation” assay, since there is no supply of food to the hatched embryos. The duration of the life stages given in the guideline is outdated and it is well-known that zebrafish embryos start feeding around 48 h after hatch and will starve without external food supply. The OECD Fish Toxicity Testing Framework report (OECD, 2012a) recommends the deletion of OECD TG 212 due to these animal welfare concerns as well as scientific concerns (low sensitivity, larval and juvenile life stages are not covered, and effects of highly lipophilic substances or specific modes of action might not be detected).

OECD TG 215, *Fish, Juvenile Growth Test* (OECD, 2000), is a chronic exposure test that assesses the effects of a chemical on the growth of juvenile fish for 28 days. At least five concentrations should be tested. The guideline does not state the number of fish per concentration / control but states that it should be based on statistical power analysis. The endpoints evaluated are weight, length, mortality, abnormal appearance, abnormal behaviour and NOEC and LOEC are determined for each effect. Recommended species are rainbow trout, zebrafish, and medaka. OECD TG 215 is rarely used since it only covers toxicity to juvenile fish.

Some regulatory frameworks recommend the use of OECD TG 212 and OECD TG 215, if it is not possible to carry out an OECD TG 210.

Requirements for fish long-term toxicity tests vary depending on the type of regulation and the geographic region (OECD, 2012a). In the EU, fish chronic toxicity data may be required for (Table 13.2):

- a) industrial chemicals (>100 t/year; EC, 2006);
- a) biocides; as part of the additional dataset, chronic fish toxicity testing may be required using OECD TG 210, 212, 215 or fish full life cycle test (OECD TG under development) (EU, 2012a),
- b) plant protection products; OECD TG 204 or OECD TG 215 over 28 days, LC50 <0.1 mg/L triggers OECD TG 210 or fish life cycle test (EC, 2009a); from 2014, revised data requirements coming into force require OECD TG 210 or fish full life cycle test (EU, 2013a; EU 2013b),
- c) veterinary pharmaceuticals; if risk quotient PEC/PNEC >1 for fish, OECD TG 210 fish early life stage test to be carried out (EMA, 2004),
- d) human pharmaceuticals; base set requirement, OECD TG 210 (EMA, 2006)

e) and others (e.g. feed, EC, 2008; EFSA, 2008).

For industrial chemicals produced at >100 t/year, Annex IX of the REACH regulation (EC, 2006) states that “*Long-term toxicity testing shall be proposed by the registrant if the chemical safety assessment according to Annex I indicates the need to investigate further the effects on aquatic organisms.*” A risk from the Chemical Safety Assessment (Annex I) is indicated if PEC/PNEC >1 and for chemicals with log K_{ow} >3 (or BCF >100) and a PEC >1/100th of the water solubility.

13.3 Standard methods to replace, reduce, or refine the use of fish

The *Threshold Approach for Acute Fish Toxicity Testing* (OECD GD 126; OECD, 2010) is a tiered testing strategy which has the potential to significantly reduce the number of fish used for acute aquatic toxicity testing. It is based on the fact that the LC₅₀/EC₅₀ value of the most sensitive of the three test species (fish, algae and invertebrates) is commonly used for hazard and risk assessment and that fish is often not the most sensitive test species.

This concept was first described for pharmaceuticals as a “threshold/step-down” approach by Hutchinson et al. (2003) and further developed for chemicals by the JRC (Jeram et al., 2005) as a threshold approach taking into consideration the requirements of the limit test as described in OECD TG 203 (Hoeger et al., 2006; ECVAM, 2006).

When fish acute toxicity data need to be generated, the fish limit test as described in OECD TG 203 (see above) should be carried out at a single concentration, the threshold concentration (TC), which corresponds to the lowest LC₅₀/EC₅₀ derived from reliable acute invertebrate (e.g. *Daphnia*, OECD TG 202) and algae (e.g. OECD TG 201) data. If no fish mortality is observed, it demonstrates that the fish is not the most sensitive test species. No further testing is required and the TC value can be used as LC₅₀ fish (LC₅₀ fish is greater than TC with 99% confidence). If mortality occurs, a full LC₅₀ test should be conducted. The Threshold Approach is recommended for acute fish toxicity testing in the biocides regulation (EU, 2012a), the REACH guidance on information requirements and chemical safety assessment (Chapter R.7b: Endpoint specific guidance; ECHA, 2012) and the OECD Fish Toxicity Testing Framework (OECD, 2012a). Based on a retrospective analysis of acute toxicity data, Creton et al. (2014) propose the use of the threshold approach for plant protection products (formulations). In fact, the recently published Commission Regulations on data requirements for active substances (2013/283/EU) and plant protection products (2013/284/EU) recommend the use of the threshold approach (EU, 2013a; EU 2013b).

At the 25th WNT (Working Group of the National Coordinators of the Test Guidelines Programme, April 2013) of the OECD, the new OECD TG 236, *Fish embryo acute toxicity (FET) test* (OECD, 2013b), was adopted and is available on the OECD website (OECD, 2013b). It is a short-term test (96 h) that determines the lethal effects of a compound on zebrafish embryos. Newly fertilised zebrafish eggs (20/concentration, water control, positive control, solvent control) are exposed to at least five concentrations of the test compound and incubated for 96 h at 26±1 °C with an appropriate light cycle (14 h light and 10 h dark). Lethal effects to the embryo (such as coagulation, absence of somite formation, non-detachment of the tail, and absence of heartbeat) are recorded on a daily basis and used for the calculation of the LC₅₀. The percentage of embryos for which at least one of the lethal observations is positive at 48 h and 96 h is plotted against tested concentrations and used for the LC₅₀ calculations.

OECD TG 236 addresses the concerns expressed over years regarding the metabolic capacity of fish embryos and the possible barrier function of the chorion (fish egg envelope). Recent studies show that zebrafish embryos have metabolic capacity (Weigt et al., 2011; Weigt et al., 2012; Kubota et al., 2011; Incardona et al., 2011), although it is not clear whether it is in the same range as that of juvenile or adult fish. OECD TG 236 recommends conducting the test with the toxic metabolites or other biotransformation products if they are believed to be more toxic than the parent compound and use this information when concluding on the toxicity of the compound.

In some cases, OECD TG 236 may not be the most appropriate test, since the overall toxic effects of a compound may not become fully evident, due to their reduced bioavailability in the test system. For example, compounds with a molecular weight $\geq 3\text{kDa}$ or a very bulky molecular structure might not pass the chorion in sufficient quantities to cause lethal effects. In fact, two polymers had been included in the OECD validation study of the zebrafish embryo acute toxicity test to challenge the barrier function of the chorion (OECD, 2012b). Neither caused consistent toxicity before hatch but toxicity significantly increased after hatch. Other compounds might cause delayed hatch and thus reduce post-hatch exposure. In these cases, other toxicity tests should be chosen.

The OECD TG 236 does not state whether the fish embryo acute toxicity test can be used as an alternative to the OECD TG 203; however, several recently published papers indicate that the LC50 values derived with the fish embryo acute toxicity test correlate well with those derived with juvenile or adult fish (Lammer et al., 2009; Knöbel et al., 2012; Belanger et al., 2013). It should be noted that at present (February 2014) the suitability of OECD TG 236 for REACH requirements (and other regulatory frameworks) still needs to be discussed and agreed upon with the regulatory authorities.

As per Article 1(3)(a)(i) of Directive 2010/63/EU (EU, 2010) on the protection of animals used for scientific purposes, live non-human vertebrate animals including independently feeding larval forms are covered by its scope. According to the description of OECD TG 236, the zebrafish embryos are used until 96 h post-fertilisation. Zebrafish is generally not considered as being capable of independent feeding until 5 days post-fertilisation. This is confirmed by the Commission Implementing Decision 2012/707/EU (EU, 2012b) on a common format on collection of information on the use of animals for scientific purposes in the EU states that "*Fish should be counted from the stage of being capable of independent feeding onward. Zebrafish kept in optimal breeding conditions (approximately + 28°C) should be counted 5 days post fertilisation*". Taking this into consideration, the zebrafish embryos in question should not be considered as "independently feeding larval forms" within the meaning of the Directive and therefore the procedure, as far as the zebrafish embryos are concerned, does not fall within its scope.

13.4 Mechanistic understanding of the endpoint

The mode of action (MOA) has been defined (ECETOC, 2007) as "*a common set of physiological and behavioural signs ... that characterise a type of adverse biological response*", whereas the mechanism of action refers to a comprehensive understanding of the entire sequence of events that result in toxicity. The report highlights the need to utilise all available information on a chemical, including mammalian toxicity data, to provide insight into a chemical MOA which will help to target testing to the most appropriate and sensitive species and identify specific tests for further characterisation. In addition, genomics,

proteomics, and biomarkers information could be used to define mechanism of action and specific biological activity and to identify critical toxicity pathway.

The definition of MOA has evolved over time and MOA is currently defined by the WHO (2009) as “A *biologically plausible sequence of key events leading to an observed effect supported by robust experimental observations and mechanistic data. A mode of action describes key cytological and biochemical events – that is, those that are both measurable and necessary to the observed effect – in a logical framework.*”

Several attempts have been made to categorise chemicals based on their mechanism or mode of action. Based on Verhaar et al. (1992), the following classification scheme was proposed:

- class 1 – non-polar narcosis or baseline toxicity; this covers inert chemicals with a non-specific mode of action (non-specific reaction between chemicals and membranes; effect depends on the hydrophobicity of the chemicals) leading *in vivo* to lethargy and narcosis;
- class 2 – polar narcosis; this covers less inert chemicals and is similar to non-polar narcosis except the fish become hyperactive before the onset of narcosis;
- class 3 – reactivity; this covers electrophilic chemicals that react in a non-specific way with the nucleophilic proteins in membranes or the cytosol and ultimately disrupt cell function and lead to death;
- class 4 –specifically acting chemicals; in this case, the targets are specific cellular receptors;
- class 5 – chemicals not assigned to class 1-4 (Enoch et al., 2008).

The Verhaar classification scheme has been implemented in the freely available Toxtree software (http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools/toxtree), as well as the OECD QSAR Toolbox (<http://www.qsartoolbox.org/>). The Toxtree implementation has been verified by Enoch et al. (2008), who also made recommendations to improve the decision tree (such as the reordering of the rules and the creation of more specific subclasses).

Russom et al. (1997) further extended the Verhaar classification scheme and distinguished 8 modes of toxic action (base-line narcosis or narcosis I, polar narcosis or narcosis II, ester narcosis or narcosis III, oxidative phosphorylation uncoupling, respiratory inhibition, electrophile/proelectrophile reactivity, acetylcholinesterase inhibition, and central nervous system seizure).

Knowledge of the different modes of action has been used to develop QSARs (see below), based on the principle that the mode of action and respective QSAR can be assigned to a chemical based on its molecular structure (McKim et al., 1987).

13.5 Status of non-standard methods and Integrated Testing Strategies

13.5.1 Quantitative Structure-Activity Relationships

There is an extensive literature on QSARs for fish toxicity, and in particular acute fish toxicity, as described in a number of reviews (Pavan et al., 2005a; Pavan et al., 2005b; Netzeva et al., 2007; 2008; Brooke et al., 2006a; Brooke et al., 2006b) and in the recently published EFSA guidance on the aquatic risk assessment of plant protection products in edge-of-field surface waters (EFSA, 2013).

Netzeva et al. (2007) summarised the features of several fish acute toxicity QSARs and found that r^2 values for most narcosis QSARs were in the region of 0.9. The evidence for the predictive power of acute fish baseline toxicity QSARs is therefore strong.

Many of these studies have focused on the development of QSARs for specific modes of action, including polar narcosis, non-polar narcosis, reactivity-based toxicity, and specifically acting mechanisms (involving non-covalent interactions with receptors or enzymes), which are the MOAs 1-4 in the Verhaar classification scheme (Verhaar et al., 1992). While it is conventional to distinguish between polar and non-polar narcosis, with the latter being slightly more toxic than the former, both modes of action can be modelled solely by the octanol-water partition coefficient, which has prompted several authors to combine the polar and non-polar narcosis MOAs into a general narcosis model (e.g. Pavan et al., 2006). It has been estimated that at least 50% of industrial chemicals act as narcotics (Pavan et al., 2005a; Pavan et al., 2005b), and for more toxic chemicals, a narcosis model provides a minimal estimate of toxicity. The so called “excess toxicity” of more toxic chemicals, i.e. the difference compared with the narcosis baseline, can often be related to specific mechanistic reaction classes (e.g. Schiff base formers, Michael acceptors, nucleophilic substitution (SN2) mechanisms) or to specific subgroups (e.g. acrylates and isothiocyanates). These reaction classes can be identified on the basis of molecular structure by using Toxtree or the OECD QSAR Toolbox.

Other studies have focused on the modelling of specific chemical classes. Examples include the QSARs implemented in the freely available ECOSAR software (<http://www.epa.gov/oppt/newchems/tools/21ecosar.htm>), and QSARs developed in the recently completed Cadaster project (Cassani et al., 2013; <http://www.cadaster.eu/>).

Relatively few QSARs have been developed for specifically acting chemicals, with the exception of some models for endocrine active chemicals (Jacobs, 2004). Another major gap concerns QSARs for chronic fish toxicity, although some have been implemented in ECOSAR (de Haas et al., 2011) and two became recently available for prediction NOEC values of MOAs1-2 chemicals (Claeys et al., 2013) and MOA1 chemicals (Austin and Eadsforth, 2014).

13.5.2 Toxicokinetic models

Stadnicka et al. (2012) investigated whether the internal concentration of chemical in fish can be predicted using a physiologically based toxicokinetic (PBTK) model and two one-compartment models. None of the models accounted for possible biotransformation and the only uptake route considered was via the gills. They compared the predicted internal concentration to the measured concentration of 39 chemicals in fish (rainbow trout, fathead minnow). The three models predicted the measured internal concentrations in fish within 1 order of magnitude for at least 68% of the chemicals; whereas, the PBTK model performed better with respect to simulating chemical concentrations in the whole body (at least 88% of internal concentrations were predicted within 1 order of magnitude using the PBTK model). The authors conclude that the three models can be used to predict concentrations in different fish species without additional experiments and propose the development of further toxicokinetic models for polar, ionisable, and easily biotransformed compounds. Moreover, Stadnicka-Michalak et al. (2014) developed a toxicokinetic model to quantify the concentration of organic chemicals in fish cells (RTgill-W1) and perform *in vitro* to *in vivo* extrapolations (IVIVE). They compared internal concentration causing effects on fish cells to *in vivo* concentration in fish gills predicted with a PBTK model and demonstrated that *in*

in vitro data can be used to determine fish internal concentrations causing long-term effects, such as reduced fish weight and length.

13.5.3 Cytotoxicity assays based on fish cell lines

These have been developed and used as research tools, and have been proposed as alternatives to the acute fish toxicity test (Castaño et al., 2003; Bols et al., 2005). More recently, they have also been used as tools to explore toxicity pathways at the molecular and cellular levels (Ankley et al., 2010). More than 280 fish cell lines have been established so far but only about 43 are listed in the international cell repositories such as the American Type Culture Collection (ATCC) and the European Collection of Cell Cultures (ECACC). After the ovary, the second most common tissue used for cultivation is the fin due to its high regenerative ability (Lakra et al., 2011). Gills and liver cells are also widely used, because of their importance in fish biology: gills are the primary target and uptake site of aquatic contaminants and are involved in gas exchange, osmoregulation and other critical functions (Lee et al., 2009); and the liver has a high metabolic capacity and detoxification function (Schirmer, 2006; Fent, 2007). For aquatic toxicity assessment, fish cell lines are preferred to mammalian cell lines as they better reflect fish-specific features, e.g. their incubation temperature is similar to the species-specific water temperature. Moreover, most of the fish cell lines are permanent, i.e. due to the spontaneously occurring immortalisation they can proliferate indefinitely. In contrast to many human cell lines, they are not cancerous, thus avoiding questions over the relevance of cancerous vs non-cancerous cells.

When using cell lines knowledge on the bioavailability of a substance *in vitro* is crucial: it can be decreased by serum components (Hestermann et al., 2000; Gulden et al., 2005) or influenced by the type of solvent used or the dosing procedure, especially for volatile and hydrophobic test chemicals (Tanneberger et al., 2010).

In the recently finalised research project, CELLsens (ECO8; <http://www.cefic-lri.org/projects>), funded by CEFIC LRI and UK-Defra, the use of a rainbow trout gill cell line-based assay (RTgill-W1) for acute fish toxicity testing was systematically evaluated taking into account various scenarios which could improve the sensitivity of fish cell lines as reviewed by Schirmer (2006). In the CELLsens project, exposure concentrations were measured at the beginning (0 h) and end of the exposure (24 h) and used to calculate the effective concentration (Tanneberger et al., 2013). The gill cell line was chosen, since gill epithelial cells are the primary site of uptake of contaminants, and damage of gill epithelial cell membrane function or other vital cell functions may affect the whole organism and result in death. In order to capture possible damage *in vitro*, several indicators for cell viability (metabolic capacity, cell membrane integrity, lysosomal membrane integrity) were measured. Moreover, the RTgill-W1 cells can sustain exposure without any serum or in a simplified buffer (L15-ex) (Schirmer et al., 1997; Ackermann and Fent, 1998).

In the CELLsens project, 35 organic chemicals were tested with the RTgill-W1 assay. Chemicals derived from a list of 60 reference chemicals (Schirmer et al., 2008) covering a wide range of *in vivo* fish toxicity, physicochemical properties and MOAs. The results (Tanneberger et al., 2013) showed a good agreement between the *in vivo* and *in vitro* effective concentrations. For up to 73% of the tested chemicals, covering chemicals acting via baseline toxicity and some with a specific MOA, the difference was less than 5-fold. For three chemicals, the fish cell assay was far less sensitive than the *in vivo* assay, two of them were neurotoxic (permethrin, lindane) and one needed metabolic activation (allyl alcohol). The outcome of the CELLsens project indicates that RTgill-W1 cytotoxicity assay is more

sensitive than other fish cell-line based assays and a promising method to reduce the use of fish for acute toxicity testing. However, further evaluation of its inter-laboratory reproducibility, predictive capacity and applicability domain (i.e. validation) is crucial prior to its use in a broader regulatory context.

Fish cell lines are also valuable tools for studying specific effects of chemicals *in vitro* and can be used to develop AOPs, chemical categories and QSAR models, or to focus toxicity testing strategies. It is however important to select the appropriate cell type and cellular or sub-cellular endpoint (Knauer et al., 2007; Tollefsen et al., 2008; Kramer et al., 2009).

13.5.4 The Adverse Outcome Pathway (AOP) concept

The so-called Adverse Outcome Pathways (AOP) concept was proposed by Ankley et al. (2010) as a means of advancing ecological risk assessment (see also Chapter 1). An AOP was defined as “a conceptual construct that portrays existing knowledge concerning the linkage between a direct molecular initiating event and an adverse outcome at a biological level of organisation relevant to risk assessment”, e.g. survival, development, reproduction at the population level for ecology. This framework incorporates the above mentioned “mechanism of action” and “mode of action” approaches. Ankley et al. (2010) provide five examples of AOPs and elaborate on their practical impact on risk assessment (the Episuite narcosis QSAR, the photoactivated toxicity of PAHs, toxicity induced via the aryl hydrocarbon receptor, activation of the oestrogen receptor and impaired vitellogenesis). AOPs can be used to focus toxicity testing in terms of species and endpoint selection, extrapolation between chemicals, and support the prediction of mixture effects (Ankley et al., 2010). Moreover, AOPs can be used as the basis for developing integrated approaches to testing and assessment (IATA; see below) and chemical categories (Schultz, 2010). An OECD guidance document on the development and assessment of AOPs has been published recently (OECD, 2013c).

In recent years, AOP-based approaches have been proposed to reduce the use of fish for long-term fish toxicity testing (e.g. with the fish early life stage test [FELS]). Two workshops organised by ILSI-HESI in 2010 and 2012 tackled this issue. Based on the discussions at the first workshop, Volz et al. (2011) proposed AOPs for 2,3,7,8-tetrachlorodibenzeno-p-dioxin (TCDD)-induced cardiotoxicity, chlorpyrifos-induced neurotoxicity, and LAS (a narcotic surfactant)-induced gill toxicity on early life stages of fish. They also proposed a three-tier testing scheme for screening and prioritising of chemicals for FELS testing, starting with high-throughput AOP-based *in vitro* screening at Tier 1, applying fish embryo test at Tier 2, and only testing chemicals positive in Tier 1 and Tier 2 in the FELS at Tier 3.

The 2nd workshop discussed which AOPs are relevant to FELS toxicity, as well as how they could be discovered and annotated (Villeneuve et al., 2014). The authors defined key events in the development of a fish embryo as development of the central nervous system, cardiovascular system, liver, kidney, etc. They gave examples of hypothesised AOPs related to impaired swim bladder inflation and reduced survival of young fish.

In 2013, the CEFIC LRI-funded project (LRI-ECO20-UA) *Development of an alternative testing strategy for the fish early life-stage test for predicting chronic toxicity* started. It aims to map FELS-relevant AOPs, develop an *in vitro* toolbox for screening FELS-relevant AOPs (Tier 1) and zebrafish embryo based assays (Tier 2).

In addition, an AOP for acute aquatic toxicity initiated by weak acid respiratory uncoupling has been proposed (OECD, 2011), as well as an AOP for paraoxon-induced toxicity during zebrafish embryogenesis (Yozzo et al., 2013).

The recent workshop on "Advancing Adverse Outcome Pathways for Integrated Toxicology and Regulatory Applications" (March 2014; Somma Lombardo Italy; co-organised by Environment Canada, US EPA, US ERDA, JRC, as well as academic institutes/organisation from USA, Switzerland, and Norway) brought together more than 50 experts from regulatory agencies, academia and industry to discuss the Adverse Outcome Pathway concept with a focus on environmental hazard and risk assessment. Reports from this workshop will become available during 2014-2015.

13.5.5 Use of fish embryos for chronic toxicity testing

Weil et al. (2009) evaluated 14 substances in the zebrafish embryo toxicity test (48h exposure) and determined their effects on lethal and sublethal effects as well as on the expression of potential marker genes (Gene-*DarT* assay). Out of seven, two marker genes (*cyp1a*, *hmx1*) appeared to be the most sensitive. For 10 substances, the LOEC derived with the Gene-*DarT* assay differed by a factor less than 10 from the LOECs derived from fish early life stage tests with zebrafish, whereas two substances were more and one substance less potent in the Gene-*DarT*.

13.5.6 Extrapolating across species and from acute to chronic effects

Regulatory aquatic risk assessment schemes require toxicity testing of chemicals on a limited number of laboratory species; thus, extrapolation from the obtained toxic responses to all species representing that trophic level in the environment is a fundamental tenet of regulatory ecotoxicological risk assessment. To derive the Predicted No-Effect Concentration (PNEC) for aquatic toxicity, safety factors are applied to the laboratory data. These factors are intended to account for interspecies differences in sensitivity, extrapolation from acute to chronic effects, and the physicochemical complexity of natural water versus laboratory test media. The choice of the factor (10, 100 or 1000) depends on the quality and quantity of the available data.

Differences in species sensitivity to acute aquatic toxicity have been well described (Weyers et al., 2000; Hutchinson et al., 2003; Jeram et al., 2005; Hoekzema et al., 2006; Tebby et al., 2011). The observation that fish is the most sensitive species in only 15-20% of the cases resulted in the development of the threshold approach (OECD GD 126; OECD, 2010), which reduces the number of fish needed for acute aquatic toxicity testing.

To address the question whether it is possible to predict acute toxicity in fish from non-vertebrate species, Netzeva et al. (2007) reviewed several Quantitative Activity-Activity Relationships (QAAR) between species, the most relevant and reliable relationship for acute fish toxicity being between *D. magna* and rainbow trout (n=360), with an r^2 value of 0.67. More recent studies have confirmed the good correlation between acute fish and daphnia toxicity data, especially for organothiophosphates ($0.74 < r^2 < 0.94$) and (benzo)triazoles ($r^2=0.87$) (Zvinavashe et al., 2009; Kar et al., 2010; Zhang et al., 2010; Cassani et al., 2013). The correlation depends on both the bio-uptake process and the MOA of the chemical (Zhang et al., 2010), as well as its physicochemical properties (Tebby et al., 2011).

The US EPA has developed a tool (ICE – Interspecies Correlation Estimation, implemented as WebICE; Raimondo et al., 2010) to predict acute toxicity to three relevant fish species (fathead minnow, rainbow trout, and common carp) on the basis of *Daphnia magna* toxicity, with the strongest correlation being evident between *D. magna* and rainbow trout.

Concerning the extrapolation from acute to chronic effects for a given substance, Ahlers et al. (2006) found that acute to chronic ratios (ACRs) vary considerably between different chemicals and can be as high as 4400. They also compared ACRs for the same substances between daphnia and fish, and found no correlation. This means that the ACR for a given substance obtained in an invertebrate species cannot be used to reliably extrapolate from the (known) acute toxicity to the (unknown) chronic toxicity in a fish species. However, there are better options for predicting chronic fish toxicity, such as the use of interspecies correlations from effects in daphnia.

13.5.7 Threshold of Toxicological Concern (TTC) approach

The TTC approach is based on the premise that there is a general exposure limit for chemicals below which no significant risk to human health or the environment is expected. By applying the TTC concept, testing on aquatic organisms can be reduced or avoided. The potential application of the TTC approach in aquatic toxicity assessment has been explored by de Wolf et al. (2005); and for endocrine active substances in the aquatic environment by Gross et al. (2010). De Wolf et al. (2005) collated reliable acute and chronic aquatic toxicity data from several ecotoxicity databases and derived an aquatic threshold of no concern (ETNC_{aq}) of 0.1 µg/L for chemicals with MOA 1-3 according to the Verhaar MOA scheme. For MOA 4 (specifically reacting) chemicals, the ETNC was significantly lower. Tolls et al. (2009) proposed using the approach for the risk assessment of poorly water soluble chemicals with MOA 1-2 and ETNC_{aq} of 1.9 µg/L. They concluded that the aquatic exposure levels of chemicals with water solubility below this ETNC_{aq} will not exceed the ecotoxicological no-effect concentration and thus their risk would be negligible.

13.5.8 Integrated testing strategies

Over the last decade, various testing strategies to reduce the use of fish in aquatic toxicity testing have been discussed in the context of the (then upcoming) REACH legislation (ECETOC, 2005; ECETOC, 2007; Grindon et al., 2008) or more recently in a broader OECD context (OECD, 2012a). The various strategies largely agree on the approach to acute aquatic toxicity testing. They start with the evaluation of existing information (chemical properties, possible aquatic exposure, existing mammalian and environmental data derived with standard or non-standard methods, etc.) and decision of its suitability for CLP and CSA. If this is not conclusive, a step-wise approach is proposed and, after each step, a decision should be taken whether the new information allows a conclusion on CLP and CSA for acute aquatic toxicity. Several possibilities to derive relevant information are listed e.g. *in silico* methods, testing on invertebrates (OECD TG 202) and algae (OECD TG 201), and use of alternatives to acute fish tests (fish cells, fish embryos, use of less fish). Some of the fish alternatives are now described in OECD documents, e.g. the threshold approach in OECD GD 126 (OECD, 2010) and the fish embryo acute toxicity test in OECD TG 236 (OECD, 2013b).

The strategies differ when it comes to long-term aquatic toxicity testing. ECETOC (2007) acknowledges that long-term testing in fish may be driven by PEC/PNEC >1 or specific regulatory requirements (e.g. human pharmaceuticals). However, there might be a low priority for fish long-term tests, if a chemical is not taken up by fish (high molecular weight

chemicals) or rapidly metabolised or excreted, i.e. having a low potential to bioaccumulate. Based on the MOA of the leading toxic effect, long-term testing would be considered and designed (MOA 3-4) or not required (MOA 1 or MOA 2). Grindon et al. (2008) recommended that long-term fish toxicity tests should only be performed if a chemical may bioaccumulate in fish. Information on potential bioaccumulation should preferably be derived by using physicochemical properties and/or *in silico* methods, and in case conclusive predictions cannot be made, a fish bioaccumulation test (OECD TG 305) should be conducted. The report of the OECD workshop “Fish toxicity testing framework” (OECD, 2012a) outlines a generic testing strategy on how to best combine OECD TGs to derive the required information. If information on aquatic long-term toxicity is required, fish toxicity testing may not be necessary, if there is evidence that daphnia or algae are more sensitive. However, if a chemical is more acutely toxic to fish or there is a risk of bioaccumulation or potential endocrine activity in fish, further testing on fish would need to be considered, e.g. the appropriate OECD TG for assessing endocrine activity or long-term fish toxicity (e.g. OECD TG 210, TG 212 or TG 215 depending on the target life stage).

A recently published review paper (Scholz et al., 2013) addresses possibilities to reduce the use of animals (birds, fish, amphibia) for environmental hazard and risk assessment of chemicals, plant protection products, biocides, pharmaceuticals and other products. Apart from reviewing currently available alternative methods, the authors discuss how animal test design and risk assessment schemes could be modified to become more resource efficient whilst providing the same level of protection. They advocate for the development of "formalised" ITS combining the various alternative approaches.

13.6 Conclusions and recommendations on aquatic toxicity

There are many options to reduce or waive the use of fish in acute aquatic toxicity testing. Some have been in use for many years, e.g. QSARs, whereas others have only recently gained acceptance, e.g. the threshold approach (OECD GD 126) and the fish embryo acute toxicity test (OECD TG 236). Both methods are mentioned in the REACH Endpoint Specific Guidance (Chapter 7b), for instance, the threshold approach is an integrated part of the ITS for CSA (where it is referred to as a limit test) and the fish embryo acute toxicity test is referred to as a possible alternative to OECD TG 203.

Since OECD TG 236 does not indicate whether the fish embryo acute toxicity test can be used as an alternative to the acute fish toxicity test, it may be useful to develop guidance on this option within the REACH legislation. Such guidance should also address the limitations of the zebrafish embryo acute toxicity test.

The results of the CellSens project show that the well-developed and standardised cytotoxicity assay based on the rainbow trout gill cell line, RTgill-W1, is a promising method and as predictive as the 48 h zebrafish embryo acute toxicity test (Tanneberger et al., 2013; Knöbel et al., 2012). However, the RTgill-W1 cytotoxicity assay is not yet formally validated. A ring trial on assessment of the transferability of RTgill-W1 cytotoxicity assay is on-going under the coordination of K. Schirmer (EAWAG, Switzerland).

QSARs provide a reliable means of predicting the acute fish toxicity of organic chemicals, especially those acting via a narcosis mode of action. Even for more reactive or specifically acting chemicals, QSARs provide estimates of the minimal (baseline) toxicity that can be expected, which may be adequate for the regulatory purpose. Relatively few QSARs are available for chronic fish toxicity, but other options can be applied; for example, the OECD

QSAR Toolbox provides extensive databases and functionalities to support grouping and read-across. The application of QSAR analysis to the modelling of fish toxicity (and aquatic toxicity in general) was very popular in the 80s and 90s, but has received relatively little attention in the 21st century. This can be attributed to the fact that the available QSARs, for narcosis at least, are as reliable as can be achieved using existing data sets, but it is also a reflection that the modelling community has shifted its efforts to other endpoints, especially those relevant to human health. Nevertheless, there is still an opportunity to develop more and improved QSARs, especially for reactive and specifically acting chemicals. Furthermore, there is merit in further investigating how different MOA classification schemes can be used in the context of TTC approaches in which specific thresholds of no environmental concern are associated with MOA classes.

There is a need to more accurately extrapolate responses across species and to identify species-specific effects. In this context, Celander et al. (2011) proposed a strategy consisting of the identification of specific MOAs / AOPs and the homology assessment of target genes, to estimate how species differences in protein functions are related to differences in *in vivo* sensitivity. In addition, knowledge on absorption, distribution, metabolism and excretion, including species differences at various life stages, will allow a more accurate extrapolation of toxic effects across species.

Several research groups are working on the identification and description of potential AOPs relevant to fish early life-stage toxicity (OECD TG 210) and CEFIC LRI launched a research project on this topic in early 2013. It is hoped that the definition of FELS-related AOPs will help to develop predictive assays using fish cell lines and fish embryos.

For the assessment of chronic fish toxicity, the options for reducing or waiving fish tests are currently more limited. However, an analysis carried out by the JRC shows that predictions can be based on toxicity data on *Daphnia* (OECD TG 211) using interspecies correlation equations (QAARs).

13.7 References

- Ackermann, G. E. and K. Fent (1998). The adaptation of the permanent fish cell lines PLHC-1 and RTG-2 to FCS-free media results in similar growth rates compared to FCS-containing conditions. *Marine Environmental Research* 46(1-5): 363-367.
- Ahlers, J., C. Riedhammer, M. Vogliano, R.-U. Ebert, R. Kühne and G. Schüürmann (2006). Acute to chronic ratios in aquatic toxicity - Variation across trophic levels and relationship with chemical structure. *Environmental Toxicology and Chemistry* 25(11): 2937-2945.
- Ankley, G. T., R. S. Bennett, R. J. Erickson, D. J. Hoff, M. W. Hornung, R. D. Johnson, D. R. Mount, J. W. Nichols, C. L. Russom, P. K. Schmieder, J. A. Serrano, J. E. Tietge and D. L. Villeneuve (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 29(3): 730-741.
- Austin, T.J. and C.V. Eadsforth (2014). Development of a chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics. *SAR QSAR Environmental Research* 25(2): 147-60.
- Belanger, S. E., J. M. Rawlings and G. J Carr. (2013). Use of fish embryo toxicity tests for the prediction of acute toxicity to chemicals. *Environmental Toxicology and Chemistry* 32(8): 1768-1783.
- Bols, N. C., V. R. Dayeh, L. E. J. Lee and K. Schirmer (2005). Chapter 2 Use of fish cell lines in the toxicology and ecotoxicology of fish. *Piscine cell lines in environmental toxicology. Biochemistry and Molecular Biology of Fishes*. T. P. Mommsen and T. W. Moon, Elsevier. Volume 6: 43-84.
- Brooke, D. and M. Crookes (2006a). Validation of (quantitative) structure-activity relationships for toxicological endpoints of regulatory importance: (Q)SARs for acute toxicity to fish - Part A. European Commission - Joint Research Centre - Institute for Health and Consumer Protection: 393 p. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_BRE_partA.pdf
- Brooke, D. and M. Crookes (2006b). Validation of (quantitative) structure-activity relationships for toxicological endpoints of regulatory importance: (Q)SARs for acute toxicity to fish - Part B. European Commission - Joint Research Centre - Institute for Health and Consumer Protection: 138 p. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_BRE_partB.pdf
- Cassani, S., S. Kovarich, E. Papa, P. P. Roy, L. van der Wal and P. Gramatica (2013). Daphnia and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity-activity modelling. *Journal of Hazardous Materials* 258-259: 50-60.
- Castaño, A., N. Bols, T. Braunbeck, P. Dierickx, M. Halder, B. Isomaa, K. Kawahara, L. Lee, C. Mothersill, P. Pärt, G. Repetto, J. Sintes, H. Ruffli, R. Smith, C. Wood and H. Segner (2003). The use of fish cells in ecotoxicology. The report and recommendations of ECVAM Workshop 47. *Alternatives to Laboratory Animals* 31(3): 317-351.
- Celander, M. C., J. V. Goldstone, N. D. Denslow, T. Iguchi, P. Kille, R. D. Meyerhoff, B. A. Smith, T. H. Hutchinson and J. R. Wheeler (2011). Species extrapolation for the 21st century. *Environmental Toxicology and Chemistry* 30(1): 52-63.
- Claeys, L., Iaccino, F., Janssen, C.R., Van Sprang, P., Verdonck, F. (2013) Development and validation of a quantitative structure-activity relationship for chronic narcosis to fish. *Environmental Toxicology and Chemistry* 32(10): 2217-25.

Creton, S., Clook, M., and J. R. Wheeler (2014). Application of the threshold approach for acute fish toxicity testing to plant protection products: a proposed framework. *Chemosphere* 96: 195-200.

de Haas, E. M., T. Eikelboom and T. Bouwman (2011). Internal and external validation of the long-term QSARs for neutral organics to fish from ECOSAR™. *SAR and QSAR in Environmental Research* 22(5-6): 545-559.

de Wolf, W., A. Siebel-Sauer, A. Lecloux, V. Koch, M. Holt, T. Feijtel, M. Comber and G. Boeije (2005). Mode of action and aquatic exposure thresholds of no concern. *Environmental Toxicology and Chemistry* 24(2): 479-485.

EC (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official Journal of the European Union* L396: 1-849.

EC (2008). Commission Regulation (EC) No 429/2008 of 25 April 2008 on detailed rules for the implementation of Regulation (EC) No 1831/2003 of the European Parliament and of the Council as regards the preparation and the presentation of applications and the assessment and the authorisation of feed additives. *Official Journal of the European Union* L133: 1-65.

EC (2009a). Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. *Official Journal of the European Union* L309: 1-47.

EC (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. *Official Journal of the European Union* L342: 59-209.

ECETOC (2005). Workshop on Alternative Testing Approaches in Environmental Risk Assessment. . ECETOC Workshop Report No 5: 34 pp.

ECETOC (2007). Intelligent testing strategies in ecotoxicology: mode of action approach for specifically acting chemicals. Technical Report No. 102: 182 pp.

ECHA (2012). REACH Guidance on information requirements and chemical safety assessment - Chapter R.7b. Endpoint specific guidance.

ECVAM (2006). ECVAM Advisory Committee Statement on the scientific validity of the upper threshold concentration (UTC) step-down approach - a new testing strategy to reduce the use of fish in acute toxicity testing. Available on: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/publication/ESAC%20statement%20UTC%20step%20down%20approach%2020060321.pdf

EFSA (2008). Technical Guidance for assessing the safety of feed additives for the environment. *The EFSA Journal* 842: 1-28.

EFSA (2013). Scientific Opinion - Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Panel on Plant Protection Products and their Residues (PPR). *EFSA Journal* 11(7): 3290.

EMA (2004). Guideline on environmental impact assessment for veterinary medicinal products Phase II, CVMP/VICH/790/03-Final (corresponds to VICH GL38).

EMA (2006). Guideline on the environmental risk assessment of medicinal products for human use (EMEA/CHMP/SWP/4447/00).

Enoch, S. J., M. Hewitt, M. T. D. Cronin, S. Azam and J. C. Madden (2008). Classification of chemicals according to mechanism of aquatic toxicity: An evaluation of the implementation of the Verhaar scheme in Toxtree. *Chemosphere* 73(3): 243-248.

EU (2010). Directive 2010/63/EU of the European Parliament and Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union L 276*: 33-79.

EU (2012a). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. *Official Journal of the European Union L167*: 1-116.

EU (2012b). Commission Implementing Decision 2012/707/EU establishing a common format for the submission of the information pursuant to Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes. *Official Journal of the European Union L 320*: 33-50.

EU (2013a). Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. *Official Journal of the European Union L93*: 1-84.

EU (2013b). Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. *Official Journal of the European Union L93*: 85-152.

Fent, K. (2007). Permanent fish cell cultures as important tools in ecotoxicology. *ALTEX:- Alternativen zu Tierexperimenten* 24 Spec No: 26-28.

Grindon, C., R. Combes, M. T. D. Cronin, D. W. Roberts and J. F. Garrod (2008). Integrated decision-tree testing strategies for environmental toxicity with respect to the requirements of the EU REACH legislation. *Alternatives to laboratory animals* 36 Suppl 1: 29-42.

Gross, M., K. Daginnus, G. Deviller, W. de Wolf, S. Dungey, C. Galli, A. Gourmelon, M. Jacobs, P. Matthiessen, C. Micheletti, E. Nestmann, M. Pavan, A. Paya-Perez, H.-T. Ratte, B. Safford, B. Sokull-Kluttgen, F. Stock, H.-C. Stolzenberg, J. Wheeler, M. Willuhn, A. Worth, J. M. Z. Comenges and M. Crane (2010). Thresholds of toxicological concern for endocrine active substances in the aquatic environment. *Integrated environmental assessment and management* 6(1): 2-11.

Gülden, M. and H. Seibert (2005). Impact of bioavailability on the correlation between *in vitro* cytotoxic and *in vivo* acute fish toxic concentrations of chemicals. *Aquatic Toxicology* 72(4): 327-337.

Hestermann, E. V., J. J. Stegeman and M. E. Hahn (2000). Cell culture serum alters the uptake and relative potencies of halogenated aromatic hydrocarbons in PLHC-1 cells. *Marine Environmental Research* 50(1-5): 545-546.

Hoeger, B., S. Jeram, M. Holt, P. Douben and M. Halder (2006). Reduction of animal use in acute aquatic toxicity testing: Further development of the threshold approach and its application to existing chemicals and plant protection products. Poster presentation at SETAC Europe 16th Annual Meeting 7-11 May 2006: abstract no MO1/AM/P05.

- Hoekzema, C. C., A. J. Murk, B. J. v. d. Waart, J. C. M. v. d. Hoeven and D. F. d. Roode (2006). Alternative approaches can greatly reduce the number of fish used for acute toxicity testing. *Environmental Toxicology and Chemistry* 25(5): 1322-1325.
- Hutchinson, T. H., S. Barret, M. Buzby, D. Constable, A. Hartmann, A. Hayes, D. Huggett, R. Länge, A. D. Lillicrap, J. O. Straub and R. S. Thompson (2003). A strategy to reduce the numbers of fish used in acute ecotoxicity testing of pharmaceuticals. *Environmental Toxicology and Chemistry* 22: 3031-3036.
- Incardona, J. P., T. L. Linbo and N. L. Scholz (2011). Cardiac toxicity of 5-ring polycyclic aromatic hydrocarbons is differentially dependent on the aryl hydrocarbon receptor 2 isoform during zebrafish development. *Toxicology and Applied Pharmacology* 257: 242-249.
- Jacobs, M. N. (2004). In silico tools to aid risk assessment of endocrine disrupting chemicals. *Toxicology* 205(1-2): 43-53.
- Jeram, S., J. M. R. Sintes, M. Halder, J. B. Fentanes, B. Sokull-Klüttgen and T. H. Hutchinson (2005). A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. *Regulatory Toxicology and Pharmacology* 42(2): 218-224.
- Kar, S. and K. Roy (2010). First report on interspecies quantitative correlation of ecotoxicity of pharmaceuticals. *Chemosphere* 81(6): 738-747.
- Knauer, K., C. Lampert and J. Gonzalez-Valero (2007). Comparison of *in vitro* and *in vivo* acute fish toxicity in relation to toxicant mode of action. *Chemosphere* 68(8): 1435-1441.
- Knöbel, M., F. J. M. Busser, A. Rico-Rico, N. I. Kramer, J. L. M. Hermens, C. Hafner, K. Tanneberger, K. Schirmer and S. Scholz (2012). Predicting adult fish acute lethality with the zebrafish embryo: Relevance of test duration, endpoints, compound properties, and exposure concentration analysis. *Environmental Science and Technology* 46(17): 9690-9700.
- Kramer, N. I., J. L. M. Hermens and K. Schirmer (2009). The influence of modes of action and physicochemical properties of chemicals on the correlation between *in vitro* and acute fish toxicity data. *Toxicology in vitro* 23(7): 1372-1379.
- Kubota, A., J. J. Stegeman, B. R. Woodin, T. Iwanaga, R. Harano, R. E. Peterson, T. Hiraga and H. Teraoka (2011). Role of zebrafish cytochrome P450 CYP1C genes in the reduced mesencephalic vein blood flow caused by activation of AHR2. *Toxicology and Applied Pharmacology* 253(21): 244-252.
- Lakra, W. S., T. R. Swaminathan and K. P. Joy (2011). Development, characterization, conservation and storage of fish cell lines: a review. *Fish Physiology and Biochemistry* 37(1): 1-20.
- Lammer, E., G. J. Carr, K. Wendler, J. M. Rawlings, S. E. Belanger and T. Braunbeck (2009). Is the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*) a potential alternative for the fish acute toxicity test? *Comparative Biochemistry and Physiology* 149(2): 196-209.
- Lee, L. E. J., V. R. Dayeh, K. Schirmer and N. Bols (2009). Applications and potential uses of fish gill cell lines: examples with RTgill-W1. *In Vitro Cellular and Developmental Biology - Animal* 45(3-4): 127-134.
- McKim, J. M., S. P. Bradbury and G. J. Niemi (1987). Fish acute toxicity syndrome and their use in the QSAR approach to hazard assessment. *Environmental Health Perspectives* 71: 171-186.

Netzeva, T. I., M. Pavan and A. Worth (2007). Review of data sources, QSARs, and integrated Testing Strategies for aquatic toxicity. European Commission Joint Research Centre Scientific and Technical Report, EUR 229443EN, Ispra, Italy.

Netzeva, T. I., M. Pavan and A. P. Worth (2008). Review of (Quantitative) Structure–Activity Relationships for Acute Aquatic Toxicity. *QSAR & Combinatorial Science* 27(1): 77-90.

OECD (1984). OECD Guidelines for Testing of Chemicals, 204. Fish, Prolonged Toxicity Test: 14-day Study. OECD, Paris, France (deleted since 2012; deletion to be implemented in 2014).

OECD (1992). OECD Guidelines for Testing of Chemicals, 203. Fish, Acute Toxicity Test. OECD, Paris, France.

OECD (1998). OECD Guidelines for Testing of Chemicals, 212. Fish, Short-term Toxicity Tests on Embryo and Sac-fry Stages. OECD, Paris, France.

OECD (2000). OECD Guidelines for Testing of Chemicals, 215. Fish, Juvenile Growth Test. OECD, Paris, France.

OECD (2010). Short Guidance on the Threshold Approach for Acute Fish Toxicity Testing. Series on Testing and Assessment No. 126; OECD, Paris, France.

OECD (2011). Report of the workshop on using mechanistic information in forming chemical categories. Series on Testing and Assessment No. 138; OECD, Paris, France.

OECD (2012a). Fish toxicity testing framework. Series on Testing and Assessment No. 171; OECD, Paris, France.

OECD (2012b). Validation report (Phase 2) for the zebrafish embryo toxicity test. Series on Testing and Assessment No. 179; OECD, Paris, France.

OECD (2013a). OECD Guidelines for Testing of Chemicals, 210. Fish, Early Life Stage Test. OECD, Paris, France.

OECD (2013b). OECD Guidelines for Testing of Chemicals, 236. Fish Embryo Acute Toxicity (FET) Test. OECD, Paris, France.

OECD (2013c). Guidance Document on developing and assessing Adverse Outcome Pathways. Series on Testing and Assessment No. 184. ENV/JM/MONO(2013)6. OECD, Paris, France.

Oris, J. T., S. E. Belanger and A. J. Bailer (2012). Baseline characteristics and statistical implications for the OECD 210 fish early-life stage chronic toxicity test. *Environmental Toxicology and Chemistry* 31(2): 370-376.

Pavan, M., A. Worth and T. I. Netzeva (2005a). Preliminary analysis of an aquatic toxicity dataset and assessment of QSAR models for narcosis. JRC Report EUR 21749 EN. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/publications

Pavan, M., A. Worth and T. I. Netzeva (2005b). Comparative assessment of QSAR models for aquatic toxicity. JRC Scientific and Technical Report EUR 21750 EN. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/publications

Pavan, M., T. I. Netzeva and A. Worth (2006). Validation of a QSAR model for acute toxicity. *SAR and QSAR in Environmental Research* 17: 147-171.

Raimondo, S., D. N. Vivian and M. G. Barron (2010). Web-based Interspecies Correlation Estimation (Web-Ice) for acute toxicity: User manual. Version 1.1. EPA/600/R-10/004 Gulf Breeze, F.L. Available from: <http://www.epa.gov/ceampubl/fchain/webice/index.html>

Rufli, H. (2012). Introduction of moribund category to OECD fish acute test and its effect on suffering and LC50 values. *Environmental Toxicology and Chemistry* 31(5): 1107-1112.

Rufli, H. and T. A. Springer (2011). Can we reduce the number of fish in the OECD acute toxicity test? *Environmental Toxicology and Chemistry* 30(4): 1006-1011.

Russom, C. L., S. P. Bradbury, S. J. Broderius, D. E. Hammermeister and R. A. Drummond (1997). Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (*Pimephales promelas*). *Environmental Toxicology and Chemistry* 16(5): 948-967.

Schirmer, K. (2006). Proposal to improve vertebrate cell cultures to establish them as substitutes for the regulatory testing of chemicals and effluents using fish. *Toxicology* 224(3): 163-183.

Schirmer, K., A. G. J. Chan, B. M. Greenberg, D. G. Dixon and N. C. Bols (1997). Methodology for demonstrating and measuring the photocytotoxicity of fluoranthene to fish cells in culture. *Toxicology in Vitro* 11(1-2): 107-119.

Schirmer, K., K. Tanneberger, N. I. Kramer, D. Völker, S. Scholz, C. Hafner, L. E. J. Lee, N. C. Bols and J. L. M. Hermens (2008). Developing a list of reference chemicals for testing alternatives to whole fish toxicity tests. *Aquatic Toxicology* 90(2): 128-137.

Scholz, S., E. Sela, L. Blaha, et al. (2013). A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regulatory Toxicology and Pharmacology* 67(3): 506-530.

Schultz, T. W. (2010). Chapter 14 Adverse Outcome Pathways: A Way of Linking Chemical Structure to In Vivo Toxicological Hazards. In *Silico Toxicology: Principles and Applications*, The Royal Society of Chemistry: 346-371.

Stadnicka, J., K. Schirmer and R. Ashauer (2012). Predicting concentrations of organic chemicals in fish by using toxicokinetic models. *Environmental Science & Technology* 46: 3273-3280.

Stadnicka-Michalak, J., Tanneberger, K., Schirmer, K., et al., 2014. Measured and modeled toxicokinetics in cultured fish cells and application to in vitro-in vivo toxicity extrapolation. *PLoS One*. 9, e92303.

Tanneberger, K., A. Rico-Rico, N. I. Kramer, F. J. M. Busser, J. L. M. Hermens and K. Schirmer (2010). Effects of Solvents and Dosing Procedure on Chemical Toxicity in Cell-Based in Vitro Assays. *Environmental Science & Technology* 44(12): 4775-4781.

Tanneberger, K., M. Knöbel, F. J. M. Busser, T. L. Sinnige, J. L. M. Hermens and K. Schirmer (2013). Predicting fish acute toxicity using a fish gill cell line-based toxicity assay. *Environmental Science & Technology* 47: 1110-1119.

Tebby, C., E. Mombelli, P. Pandard and A. R. R. Péry (2011). Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish. *Science of the Total Environment* 409(18): 3334-3343.

Tollefsen, K. E., C. Blikstad, S. Eikvar, E. Farmen Finne and I. Katharina Gregersen (2008). Cytotoxicity of alkylphenols and alkylated non-phenolics in a primary culture of rainbow trout (*Onchorhynchus mykiss*) hepatocytes. *Ecotoxicology and Environmental Safety* 69(1): 64-73.

Tolls, J., M. Muller, A. Willing and J. Steber (2009). A new concept for the environmental risk assessment of poorly water soluble compounds and its application to consumer products. *Integrated Environmental Assessment and Management* 5(3): 374-378.

Verhaar, H. J. M., C. J. van Leeuwen and J. L. M. Hermens (1992). Classifying environmental pollutants. *Chemosphere* 25(4): 471-491.

Villeneuve, D., Volz, D.C., Embry, M.R., Ankley, G., Belanger, S.E., Léonard, M., Schirmer, K., Tanguay, R., Truong, L, and Wehmas, L. (2014) Investigating alternatives to the fish early-life stage test: A strategy for discovering and annotating adverse outcome pathways for early fish development. *Environmental Toxicology and Chemistry* 33(1): 158-169.

Volz, D. C., S. Belanger, M. Embry, S. Padilla, H. Sanderson, K. Schirmer, S. Scholz and D. Villeneuve (2011). Adverse Outcome Pathways during Early Fish Development: A Conceptual Framework for Identification of Chemical Screening and Prioritization Strategies. *Toxicological Sciences* 123(2): 349-358.

Weigt, S., N. Hübler, N. Strecker, T. Braunbeck and T. H. Broschard (2011). Zebrafish (*Danio rerio*) embryos as a model for testing proteratogens. *Toxicology* 281: 25-36.

Weigt, S., N. Hübler, R. Strecker, T. Braunbeck and T. H. Broschard (2012). Developmental effects of coumarin and the anticoagulant coumarin derivative warfarin on zebrafish (*Danio rerio*) embryos. *Reproductive Toxicology* 33: 133-141.

Weil, M., S. Scholz, M. Zimmer, F. Sacher and K. Duis (2009). Gene expression analysis in zebrafish embryos: a potential approach to predict effect concentrations in the fish early life stage test. *Environmental Toxicology and Chemistry* 28(9): 1970-1978.

Weyers, A., B. Sokull-Klüttgen, J. Baraibar-Fentanes and G. Vollmer (2000). Acute toxicity data: A comprehensive comparison of results of fish, daphnia, and algae tests with new substances notified in the European Union. *Environmental Toxicology and Chemistry* 19(7): 1931-1933.

WHO (2009). Principles and Methods for the Risk Assessment of Chemicals in Food. *Environmental Health Criteria* 240. WHO, Geneva, Switzerland (Definitions page A-25). Available from: <http://www.who.int/entity/foodsafety/chem/principles/en/index.html>

Yozzo, K. L., S. P. McGee and D. C. Volz (2013). Adverse outcome pathways during zebrafish embryogenesis: A case study with paraoxon. *Aquatic Toxicology* 126(0): 346-354.

Zhang, X. J., H. W. Qin, L. M. Su, W. C. Qin, M. Y. Zou, L. X. Sheng, Y. H. Zhao and M. H. Abraham (2010). Interspecies correlations of toxicity to eight aquatic organisms: Theoretical considerations. *Science of the Total Environment* 408(20): 4549-4555.

Zvinavashe, E., T. Du, T. Griff, H. H. J. v. d. Berg, A. E. M. F. Soffers, J. Vervoort, A. J. Murk and I. M. C. M. Rietjens (2009). Quantitative structure-activity relationship modeling of the toxicity of organothiophosphate pesticides to *Daphnia magna* and *Cyprinus carpio*. *Chemosphere* 75(11): 1531-1538.

Table 13.1. OECD test guidelines for fish toxicity testing

OECD Guideline	203 – Fish, Acute Toxicity Test	236 – Fish Embryo Acute Toxicity (FET) Test	204 – Fish, Prolonged Toxicity Test	210 - Fish, Early-life Stage Toxicity Test (revised in 2013)	212 - Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages	215 – Fish, Juvenile Growth Test
Species	zebrafish (<i>Danio rerio</i>), fathead minnow (<i>Pimephales promelas</i>), Japanese medaka (<i>Oryzias latipes</i>), rainbow trout (<i>Oncorhynchus mykiss</i>) bluegill sunfish ¹ (<i>Lepomis macrochirus</i>), common carp ¹ (<i>Cyprinus carpio</i>), guppy ¹ (<i>Poecilia reticulata</i>)	zebrafish (<i>Danio rerio</i>); in development for fathead minnow	See OECD TG 203	rainbow trout (<i>Oncorhynchus mykiss</i>) zebrafish (<i>Danio rerio</i>), fathead minnow (<i>Pimephales promelas</i>), Japanese medaka (<i>Oryzias latipes</i>) sheepshead minnow ² (<i>Cyprinodon variegatus</i>), silverside ² (<i>Menidia</i> sp)	zebrafish (<i>Danio rerio</i>), Japanese medaka (<i>Oryzias latipes</i>), rainbow trout (<i>Oncorhynchus mykiss</i>), fathead minnow (<i>Pimephales promelas</i>), common carp (<i>Cyprinus carpio</i>), goldfish (<i>Carassius auratus</i>) bluegill (<i>Lepomis macrochirus</i>) tidewater silverside ² (<i>Menidia peninsulae</i>) herring ² (<i>Clupea harengus</i>) cod ² (<i>Gadus morhua</i>) sheepshead minnow ² (<i>Cyprinodon variegatus</i>)	rainbow trout (<i>Oncorhynchus mykiss</i>) is recommended species zebrafish (<i>Danio rerio</i>), Japanese medaka (<i>Oryzias latipes</i>)
Life-stages covered	Juvenile or adult	fertilised eggs, embryos until 96h post-fertilisation	Juvenile or adult	fertilised eggs, embryo, sac-fry, larvae juvenile fish	fertilised eggs, embryo, sac-fry	juvenile
Number of animals (min)	42 (at least 7-10 per	120 (at least 20 fertilised	10 per concentration & control	480 (at least	150 (at least 30 fertilised	Not defined; depends on the test design; at

OECD Guideline	203 – Fish, Acute Toxicity Test	236 – Fish Embryo Acute Toxicity (FET) Test	204 – Fish, Prolonged Toxicity Test	210 - Fish, Early-life Stage Toxicity Test (revised in 2013)	212 - Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages	215 – Fish, Juvenile Growth Test
	concentrations & control)	eggs per concentration & control) ³		80/concentration & control)	eggs per concentration & control)	least two replicates per concentration
Concentrations	At least 5	At least 5	Not stated; should allow calculation of LC50, ECx, NOEC	At least 5 (or less when only NOEC needed; see limit test)	At least 5 (or less with justification)	At least 5
Controls	Water, and if needed solvent	Water, and if needed solvent	Water, and if needed solvent	Water, and if needed solvent	Water, and if needed solvent	Water, and if needed solvent
Duration	96 h	96 h	14 days	28-60 days (species dependant)	Until embryos start feeding or onset of mortality due to starvation (for zebrafish TG states 8-10 days)	28 days
Endpoint	Lethal effects LC50	Lethal effects LC50	Lethal & sublethal effects (appearance, size, behaviour) LC50, ECx, NOEC	<i>At all life stages:</i> hatching success, abnormal appearance, abnormal behaviour, survival / mortality <i>At the end of the test</i> wet weight, length NOEC, LOEC (or ECx) for each observation	<i>Embryo/sac-fry stage:</i> hatching success, abnormal appearance, abnormal behaviour, survival / mortality <i>At the end of the test:</i> weight, length	Growth rate as ECx; NOEC, LOEC
Effects in control	Max 1 control fish can die	Survival rate >90%	Effects <10%	Thresholds for hatching success, post-hatch survival (depending on the species)	Thresholds for hatching success, post-hatch survival (depending on the species)	<10% mortality; >50% weight gain
Limit test	Limit test at single concentration >100 mg/L; control(s); 7-10 fish	Limit test at single concentration >100 mg/L; control(s); at least 20 fertilised eggs	-	Limit test; concentrations above LC50 derived with TG 203 or >10 mg/L not to be tested	Limit test; concentrations above LC50 derived with TG 203 or >100 mg/L not to be tested	-

1 = warm water species; 2 = estuarine or saltwater species; 3 = Depending on regulatory national framework, may not be considered an animal test
LC = Lethal Concentration; NOEC = No Observed Effect Concentration; EC_x = x% Effect Concentration; LOEC = Lowest Observed Effect Concentration; OECD TG 204
deleted in 2012

Table 13.2. Fish toxicity tests – Regulatory requirements across sectors

Regulatory framework	Endpoint		
	Short-term fish toxicity	Long-term fish toxicity	Bioaccumulation
REACH	<p>>10 t/year</p> <p>Not to be conducted if:</p> <ul style="list-style-type: none"> - aquatic toxicity is unlikely to occur, e.g. chemical highly insoluble in water or unlikely to cross biological membranes or - a long-term aquatic toxicity study on fish is available <p>In general applicants are invited to consider long-term aquatic toxicity testing:</p> <ul style="list-style-type: none"> - as described in Annex IX <ul style="list-style-type: none"> o shall be considered if the chemical safety assessment according to Annex I indicates the need to investigate further effects on aquatic organisms. o if the substance is poorly water soluble. 	<p>>100 t/year</p> <ul style="list-style-type: none"> - Long-term toxicity testing shall be proposed by the registrant if the chemical safety assessment according to Annex I indicates the need to investigate further the effects on aquatic organisms. The choice of the appropriate test(s) depends on the results of the chemical safety assessment. - Appropriate tests: Fish early-life stage (FELS) toxicity test (OECD 210); Fish short-term toxicity test on embryo and sac-fry stages (OECD 212) or Fish, juvenile growth test (OECD 215) 	<p>>100 t/year</p> <p>Not to be conducted if:</p> <ul style="list-style-type: none"> - the substance has a low potential for bioaccumulation (for instance a log $K_{ow} \leq 3$) and/or a low potential to cross biological membranes, or - direct and indirect exposure of the aquatic compartment is unlikely.
Plant protection products (<i>until 2013</i>) (EC, 2009a)	<p><i>Active substance:</i> OECD TG 203 (rainbow trout and warm water species)</p> <p><i>Formulations:</i> OECD TG 203 if fish is the most sensitive species for the active substances included</p>	<p><i>Active substance:</i> OECD TG 204 or OECD TG 215 over 28 days, LC50 <0.1 mg/L triggers TG 210 or fish life cycle test</p>	<p><i>Active substance</i> To be conducted if</p> <ul style="list-style-type: none"> - log $K_{ow} > 3$ or other indications that the chemical may bioaccumulate in fish, - the substance is stable in water <p><i>Formulations:</i> fish bioconcentration studies are triggered when log $P_{ow} > 3$ and $dt_{90} (w/s) > 10$ d or if >1 application</p>
Plant protection products data requirements (<i>from 2014 onwards</i>) (EU, 2013a; 2013b)	<p><i>Active substance:</i> Threshold approach or OECD TG 203 (rainbow trout)</p>	<p><i>Active substance:</i> Fish early-life stage (FELS) toxicity test (OECD TG 210)</p> <ul style="list-style-type: none"> - if exposure of surface water likely and the substance stable in water; <p>Fish full life cycle test may be required</p>	<p><i>Active substance:</i> To be conducted if</p> <ul style="list-style-type: none"> - log $K_{ow} > 3$ or other indications that the chemical may bioaccumulate in fish, - the substance is stable in water

Regulatory framework	Endpoint		
	Short-term fish toxicity	Long-term fish toxicity	Bioaccumulation
	<p><i>Product:</i> Only if the toxicity cannot be predicted from the active substance(s) or read-across</p> <p>In case that fish data are needed: Threshold approach or OECD TG203 (rainbow trout)</p>	<p><i>Product:</i> Only if the toxicity cannot be predicted from the active substance(s)</p> <p>Necessary studies should be discussed with competent authority.</p>	
Biocides (EU, 2012a)	<p><i>Core data set:</i> Short-term toxicity on fish (use of threshold approach). If a valid long-term fish toxicity study is available, the short-term fish toxicity study can be waived</p>	<p><i>Additional data set:</i> depending on the results of the studies on fate / behaviour and intended use, e.g. risk to the aquatic environment, and long-term exposure - long-term toxicity studies on fish may be required, one or two tests, OECD 210, 212, 215 or fish full life cycle test; bioaccumulation in an appropriate aquatic species.</p>	<p><i>Core data set:</i> Experimental bioconcentration test can be waived if log Kow <3 or other evidence that the substance has a low potential for bioconcentration</p> <p><i>Additional data set:</i> depending on the results of the studies on fate / behaviour and intended use, e.g. risk to the aquatic environment, bioaccumulation in an appropriate aquatic species may be required</p>
Veterinary Pharmaceuticals (EMA, 2004) Stepwise approach with initial screening (Phase I) to identify exposure, bioaccumulation, persistence. If given then studies are performed (Phase II):	Tier A – base set requirement OECD TG 203	if risk quotient (PEC/PNEC) >1 for fish, OECD TG 210 fish early life stage test	Tier B – if log Kow >4 and evidence for bioaccumulation from other studies, OECD 305 to be carried out
Human Pharmaceuticals (EMA, 2006) Stepwise approach with initial screening (Phase I) to identify exposure, bioaccumulation, persistence. If given then studies are performed (Phase II):	-	Tier A – base set requirement OECD TG 210 – fish early life stage test	Tier B – depending on information on fate in Tier A, bioconcentration study with drug substance or its metabolites
Feed (EC, 2008; EFSA, 2008) Stepwise approach with initial determination (Phase I) whether a	Phase IIa: OECD TG 203	Phase IIb: OECD TG 210	

Regulatory framework	Endpoint		
	<i>Short-term fish toxicity</i>	<i>Long-term fish toxicity</i>	<i>Bioaccumulation</i>
significant environmental effect of the additive is likely (based on estimated PEC). If likely, then studies are performed (Phase II):			
Cosmetic ingredients (EC, 2009b)	See REACH	See REACH	See REACH

Chapter 13. Table of References with Notes

Traditional fish tests – OECD TGs and GD	
<i>Short-term fish toxicity</i>	
OECD (1992a). Guidelines for Testing of Chemicals, 203, Fish, acute toxicity test. Paris, France. Available from: http://www.oecd.org	The fish acute toxicity test determines the concentration of a compound which is lethal to 50% of the fish (LC50) at 24, 48, 72, 96 h exposure. At least 5 concentrations; groups of 7-10 fish; 7-10 control (dilution water) fish and if used also solvent control group with 7-10 fish; LC0 and LC100 determined (latter not used in risk assessment). Allows limit test at a single concentration (= 100 mg/L) with 7-10 fish in the treatment group and 7-10 fish in the control group / solvent group. If no mortality occurs LC50 is reported as >100mg/L. If mortality occurs, full test is required. Species: zebrafish, fathead minnow, medaka, rainbow trout, bluegill sunfish (and others)
Rufli and Springer (2011)	<i>See below Status of Non-standard methods // Reduction / refinement general</i>
Rufli (2012)	<i>See below Status of Non-standard methods // Reduction / refinement general</i>
<i>Note: TG 204 will be deleted in April 2014</i>	
OECD (1984). Guidelines for Testing of Chemicals, 204. Fish, Prolonged Toxicity Test: 14-day Study. OECD, Paris, France (deleted since 2012; deletion to be implemented in 2014)	Determines threshold levels for lethal effects, observed (non-lethal) effects and NOEC after 14 day exposure (option to extend by 2 weeks); 10 fish per concentration, control group, solvent group if necessary; number of concentrations not defined, should be chosen to allow determining threshold levels for the lethal and other observable effects, NOEC; not necessary to test beyond 100 mg/L if threshold levels not reached. Species: as TG 203 OECD Fish Toxicity Testing Framework report recommends deletion of TG 204, since it is hardly performed and not well defined. Moreover, OECD TG 203 allows extension of exposure beyond 96 h, in case slow mortality occurs.
OECD (2010). Series on Testing and Assessment No. 126; Short Guidance on the Threshold Approach for Acute Fish Toxicity Testing. OECD, Paris, France. Available from: http://www.oecd.org	The threshold approach for acute fish toxicity testing is a reduction strategy and based on the observation that fish are not always the most sensitive of the three test species used in acute aquatic toxicity testing. Testing of fish is performed at a single concentration (= threshold concentration, TC) following the limit test as described in TG 203. The threshold concentration is derived from reliable invertebrate (e.g. daphnia) and algae data and corresponds to the lowest of the LC50/EC50 value.
OECD (2013b). Guidelines for Testing of Chemicals, 236. Fish Embryo Acute Toxicity (FET) Test. OECD, Paris, France. Available from: http://www.oecd.org	The fish embryo acute toxicity test determines the concentration of a compound which is lethal to 50% of the embryos at 48 h and 96 h exposure. At least 5 concentrations; 20 fertilised eggs / concentration, positive control (3,4-dichloroaniline), water control, solvent control if required; the following observations are recorded and indicate the death of the fish embryo: coagulation, absence of somites (no muscular structure, embryo will not be able to hatch and will die), non-detachment of the tail (functional tail is needed for the hatching),

	<p>and absence of heartbeat. As TG 203, TG 236 allows a limit test at a single concentration (= 100mg/L) with 20 embryos in the treatment group and 20 embryos in the respective control groups. If no mortality occurs, the LC50 is reported >100mg/L. If mortality occurs, full test is required. If a compound needs to get activated via metabolism and the metabolites or transformation products are believed to be more toxic than the parent compound, the test should also be carried out with the relevant metabolites / transformation products and the results be considered when concluding on the toxicity. The test might not be appropriate for compounds with a molecular weight ≥ 3kDa, a very bulky molecular or compounds precluding or delaying hatch, since in these cases zebrafish embryos will not be fully exposed to the compound and due to decreased bioavailability of the compound toxic effect may not develop.</p>
<p>Long-term fish toxicity</p> <p>OECD (1992). Guidelines for Testing of Chemicals, 210. Fish, Early-life Stage Toxicity Test. (out of date)</p> <p>OECD (2013a). Guidelines for Testing of Chemicals, 210. Fish, Early Life Stage Test. OECD, Paris, France. Available at: http://www.oecd.org</p>	<p>As stated before the revision: Chronic exposure. Assesses effects (lethal, sublethal) of a chemical on early life stages of fish (embryos, larvae, juvenile fish), endpoints are: hatching, abnormal appearance, abnormal behaviour, survival / mortality at embryo, larval, juvenile stage and overall; dry weight and length at the end of the test; NOEC and LOEC (not always possible) for each endpoint. Starts fertilised eggs and continues until the controls are freely feeding. At least 5 concentrations with at least 60 eggs / concentration divided in two replicates; control group and solvent control group if needed; Species: freshwater - zebrafish, fathead minnow, medaka, rainbow trout; saltwater sheephead minnow</p> <p>Revised guideline: main differences At least 80 eggs / concentration / control; divided into at least four replicates; wet weight instead of dry weight; performance standards set for control fish (hatching success, post-hatch survival); guidance on statistical analysis for NOEC determination and ECx determination; allows conducting a limit test or using less than five concentrations when only an empirical NOEC is needed.</p>
<p>OECD (1998). Guidelines for Testing of Chemicals, 212. Fish, Short-term Toxicity Tests on Embryo and Sac-fry Stages. OECD, Paris, France. Available from: http://www.oecd.org</p>	<p>Sub-chronic exposure: exposure starts with eggs and should end before the yolk-sac is completely absorbed or before mortality due to starvation begins. Assesses effects of a chemical on hatching, abnormal appearance, abnormal behaviour, survival / mortality at embryo and larval stages and overall; weight and length at the end of the test; NOEC and LOEC. May be used as range finding test for OECD TG 210. Less sensitive compared to OECD TG 210, in particular for highly lipophilic substances or substances with specific</p>

	<p>mode of action. Species: zebrafish, fathead minnow, medaka, rainbow trout, carp.</p> <p>Note: There is a major animal welfare concern associated with OECD TG 212 being also called the fish "starvation" assay, since no food is provided to the hatched embryos during the test. Many species already start feeding at the sac-fry stage, e.g. zebrafish embryo start feeding around 48 h after hatch. OECD Fish Toxicity Testing Framework report (OECD, 2012) recommends its deletion due to the animal welfare concerns and its low sensitivity.</p>
<p>OECD (2000). Guideline for Testing of Chemicals, 215. Fish, Juvenile Growth Test. OECD, Paris, France. Available from: http://www.oecd.org</p>	<p>Chronic exposure: assesses effects of a chemical on the growth of juvenile fish for 28 days; at least five concentrations; number of fish per concentration/control not stated but should be based on statistical power. Effects: weight, length, mortality, abnormal appearance, abnormal behaviour; NOEC and LOEC for each effect</p> <p>Species: rainbow trout, zebrafish, medaka</p> <p>Note: only rarely used; does not cover all life stages</p>
<p><i>Bioaccumulation / bioconcentration</i></p>	
<p>OECD (2012). Guidelines for Testing of Chemicals, 305, Bioaccumulation in Fish: Aqueous and Dietary Exposure. Paris, France. Available from: http://www.oecd.org</p>	<p>OECD TG 305 has recently been revised and now includes two exposure routes: aqueous and dietary. The latter is recommended for strongly hydrophobic substances (log Kow > 5 and solubility below ~ 0.01-0.1 mg/L).</p> <p><i>305-I</i> Aqueous exposure: BCF is determined by exposure of fish for (usually) 28 days (uptake phase) followed by the depuration phase (usually half of the duration of the uptake phase). Fish are samples (4 x sampling) on at least five occasions during the uptake and at least four occasions during the depuration phase. At least 108 fish are used, i.e. 36 per concentration or control.</p> <p>BCF should be normalised for lipid concentration and fish weight. Possible to use only one concentration for non-polar narcotics.</p> <p><i>305-II</i> Minimised aqueous exposure: to be used only for non-polar organic substances; sampling is reduced to four; thus less fish are used, i.e. 48 for a study with two concentrations and water control.</p> <p><i>305-III</i>: Dietary Exposure Bioaccumulation Fish Test</p> <p>Should be used for substances where the aqueous exposure methodology is not practicable (for example because stable, measurable water concentrations cannot be maintained, or adequate body burdens cannot be achieved within 60 days of exposure); the endpoint from this test will be a dietary biomagnification factor (BMF). The test consists of two phases: uptake (test substance-spiked feed) and depuration (clean, untreated feed). Depending on the test design and sampling scheme, at least 50-120 fish (treatment group), 50-110 control fish, and if needed 15 fish for lipid correction are used.</p>
<p>Regulatory requirements across sectors</p>	

<p>EC (2006). Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L396, 1-849.</p>	<p><i>Chemicals:</i> Information requirements depend on the production volume of the chemical concerned:</p> <ul style="list-style-type: none"> • Annex VIII (10-100 tpa): short-term fish toxicity <ul style="list-style-type: none"> ○ The study does not need to be conducted <ul style="list-style-type: none"> ▪ there are mitigating factors indicating that aquatic toxicity is unlikely to occur, for instance if the substance is highly insoluble in water or the substance is unlikely to cross biological membranes or ▪ if a long-term aquatic toxicity study on fish is available ○ in general applicants are invited to consider long-term aquatic toxicity testing ○ long-term toxicity testing as described in Annex IX <ul style="list-style-type: none"> ▪ shall be considered if the chemical safety assessment according to Annex I indicates the need to investigate further effects on aquatic organisms. ▪ if the substance is poorly water soluble. • Annex IX-X (100->1000 tpa): <ul style="list-style-type: none"> ○ long-term toxicity testing: <ul style="list-style-type: none"> • Long-term toxicity testing shall be proposed by the registrant if the chemical safety assessment according to Annex I indicates the need to investigate further the effects on aquatic organisms. The choice of the appropriate test(s) depends on the results of the chemical safety assessment. • Appropriate tests: Fish early-life stage (FELS) toxicity test (OECD TG 210); Fish short-term toxicity test on embryo and sac-fry stages (OECD TG 212) or Fish, juvenile growth test (OECD TG 215) ○ Bioaccumulation in aquatic species, preferably fish <ul style="list-style-type: none"> • The study need not be conducted if: <ul style="list-style-type: none"> • the substance has a low potential for bioaccumulation (for instance a $\log K_{ow} \leq 3$) and/or a low • potential to cross biological membranes, or • direct and indirect exposure of the aquatic compartment is unlikely.
<p>EC (2009a). Regulation (EC) No 1107/2009 of the European</p>	<p><i>Plant protection products:</i> Requirements are stated in <i>SANCO/3268/2001 rev.4 (final) 17</i></p>

<p>Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309: 1-47.</p> <p><i>From 2014 onwards</i> EU (2013a). Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union L93: 1-84.</p> <p>EU (2013b). Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union L93: 85-152.</p>	<p><i>October 2002, Working Document, Guidance document on Aquatic Ecotoxicology in the context of the Directive 91/414/EEC – (from 2014 onwards obsolete; superseded by EU 2013a and EU2013b)</i></p> <p><i>Active substance:</i> Short-term fish toxicity (OECD TG 203) with rainbow trout and one warm water species; long-term/chronic fish toxicity: OECD TG 204 or OECD TG 215 over 28 days, LC50 <0.1 mg/L triggers TG 210 or fish life cycle test</p> <p><i>Formulations:</i> Acute fish toxicity test if fish is the most sensitive species for the active substances included; fish bioconcentration studies are triggered when log Pow > 3 and dt90 (w/s) > 10 d or if > 1 application</p> <p><i>Active substance:</i> Short-term fish toxicity (threshold approach (OECD GD126) or OECD TG 203 with rainbow trout); long-term fish toxicity with OECD TG 210 if exposure of surface water likely and the substance stable in water; fish full life cycle test may be required; Bioconcentration (fish) studies are required if log Kow > 3 or other indications that the chemical may bioaccumulate in fish, if the substance is stable in water</p> <p><i>Products:</i> short-term fish toxicity only if the toxicity cannot be predicted from the active substance(s) or read-across. In case that fish data are needed: Threshold approach or OECD TG203 (rainbow trout); long-term toxicity: Only if the toxicity cannot be predicted from the active substance(s). Necessary studies should be discussed with competent authority.</p>
<p>EU (2012a). Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. Official Journal of the European Union L167: 1-116.</p>	<p><i>Biocides:</i> requirements are outlined in Annex II.</p> <p>Core data set: short-term toxicity on fish (use of threshold approach; OECD GD126). If a valid long-term fish toxicity study is available, the short-term fish toxicity study can be waived.</p> <p>Experimental bioconcentration test can be waived if log Kow <3 or other evidence that the substance has a low potential for bioconcentration.</p> <p>Additional data set: depending on the results of the studies on fate / behaviour and intended use, e.g. risk to the aquatic environment, and long-term exposure - long-term toxicity studies on fish may be required, one or two tests, OECD TG 210, TG 212, TG 215 or fish full life cycle test; bioaccumulation in an appropriate aquatic species.</p>
<p>EMA (2004). Guideline on environmental impact assessment for</p>	<p><i>Pharmaceuticals for veterinary use:</i> Stepwise approach with initial screening (Phase I) to</p>

<p>veterinary medicinal products Phase II, CVMP/VICH/790/03-Final (corresponds to VICH GL38). [based on outcome of Guideline on environmental impact assessment for veterinary medicinal products – Phase I (CVMP/VICH/592/98 // VICH GL6)]</p>	<p>identify exposure, bioaccumulation, persistence. If given then studies are performed (Phase II): Tier A – base set requirement OECD TG 203 – fish acute toxicity test; Tier B – if log Kow>4 and evidence for bioaccumulation from other studies, OECD TG 305 to be carried out; if risk quotient (PEC/PNEC) > 1 for fish, OECD TG 210 fish early life stage test (OECD TG 212, TG 215 not recommended)</p>
<p>EMA (2006). Guideline on the environmental risk assessment of medicinal products for human use (EMEA/CHMP/SWP/4447/00).</p>	<p><i>Pharmaceuticals for human use:</i> Stepwise approach with initial screening (Phase I) to identify exposure, bioaccumulation, persistence. If given then studies are performed (Phase II): Tier A – base set requirement OECD TG 210 – fish early life stage test; Tier B – depending on information on fate in Tier A, bioconcentration study with drug substance or its metabolites</p>
<p>EC 2008 Commission Regulation (EC) No 429/2008 of 25 April 2008 on detailed rules for the implementation of Regulation (EC) No 1831/2003 of the European Parliament and of the Council as regards the preparation and the presentation of applications and the assessment and the authorisation of feed additives. Official Journal of the European Union L133: 1-65</p> <p>EFSA (2008). Technical Guidance for assessing the safety of feed additives for the environment The EFSA Journal 842: 1-28.</p>	<p><i>Feed additives:</i> Stepwise approach with initial determination (Phase I) whether a significant environmental effect of the additive is likely (based on estimated PEC). If likely, acute fish toxicity (OECD TG 203) would be tested in Phase IIa and chronic fish toxicity (OECD TG 210) in Phase IIb.</p>
<p>EC (2009b). Regulation (EC) No 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. Official Journal of the European Union L342: 59-209.</p>	<p><i>Cosmetics:</i> The environmental concerns that substances used in cosmetic products may raise are considered through the application of Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and establishing a European Chemicals Agency(4) OJ L 396, 30.12.2006, p. 1. (4), which enables the assessment of environmental safety in a cross-sectoral manner.</p>
<p>EC (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353: 1- 1355. EC (2008).</p>	<p><i>Art 7: Where new tests are carried out for the purposes of this Regulation, tests on animals within the meaning of Directive 86/609/EEC shall be undertaken only where no other alternatives, which provide adequate reliability and quality of data, are possible.</i> Four classification categories – Acute Category 1 and Chronic Categories 1 – 3; Safety net classification = Chronic Category 4; considered are acute aquatic toxicity (fish, daphnia, algae), potential for or actual bioaccumulation, degradation, chronic aquatic toxicity (fish, daphnia, algae)</p>

Standard methods to replace, reduce or refine the use of fish	
<i>Threshold Approach for fish acute toxicity testing</i>	
OECD GD126 (OECD, 2010)	<i>See above – Acute fish toxicity guidelines</i>
Hutchinson et al. (2003)	<i>See below REDUCTION - Step-down Approach</i>
Jeram et al. (2005)	<i>See below REDUCTION - Step-down Approach</i>
Hoeger, B., S. Jeram, M. Holt, P. Douben and M. Halder (2006). Reduction of animal use in acute aquatic toxicity testing: Further development of the threshold approach and its application to existing chemicals and plant protection products. Poster presentation at SETAC Europe 16th Annual Meeting 7-11 May 2006: abstract no MO1/AM/P05.	Both, the step-down approach (see below) and the threshold approach (OECD GD126) start with the derivation of the (upper) threshold concentration = the lowest of EC/LC 50 of daphnia / algae tests. An acute fish test is carried out at this concentration using 5 test / 5 control fish (step-down approach) or 7-10 test / 7-10 control fish as given in the limit test according to OECD TG 203. When mortality occurs, the threshold approach requires a full OECD TG 203 to be conducted; whereas in the step-down approach testing is continued at lower concentrations until no mortality occurs. The LC50 is then calculated based on the mortality rates of the individual tests. The ECVAM Scientific Advisory Committee only endorsed the threshold approach. ECVAM applied the threshold (step-down) and threshold approach to various datasets (fish, daphnia, algae) retrieved for plant protection products and industrial chemicals from various databases (IUCLID4, ECETOC, DG SANCO) databases. Depending on the database, fish was the most sensitive species for 17-30% chemicals/plant protection products and potential for reduction was between 32 % for the threshold approach and up to 60% for the threshold (step-down) approach.
ECVAM (2006). ECVAM Advisory Committee Statement on the scientific validity of the upper threshold concentration (UTC) step-down approach - a new testing strategy to reduce the use of fish in acute toxicity testing. Available on: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/publication/ESAC%20statement%20UTC%20step%20down%20approach%2020060321.pdf	The ECVAM Scientific Advisory Committee evaluated the threshold/step-down approach as proposed by Jeram et al. (2005) and other scenarios proposed in the ESAC dossier. ESAC approved the threshold approach. For further explanation see OECD GD126 and Hoeger et al., 2006 (above).
<i>Zebrafish embryo toxicity test</i>	
OECD TG 236 (OECD 2013b)	<i>See above – Acute fish toxicity guidelines</i>
Weigt, S., N. Hübler, N. Strecker, T. Braunbeck and T. H. Broschard (2011). Zebrafish (<i>Danio rerio</i>) embryos as a model for testing proteratogens. Toxicology 281: 25-36.	The authors tested 10 well-know proteratogens in a zebrafish embryo toxicity test (72 h exposure; specific malformation endpoints; proposed as teratogenicity screening test for human health to reduce reproductive toxicity studies) and found that 9 caused teratogenic effects in zebrafish embryos. They conclude from these results that zebrafish embryos can bioactivate the proteratogens without any addition of exogenous metabolic activation systems and have phase I enzyme activity (CYP) at very early stages of development.

<p>Weigt, S., N. Hübner, R. Strecker, T. Braunbeck and T. H. Broschard (2012). Developmental effects of coumarin and the anticoagulant coumarin derivative warfarin on zebrafish (<i>Danio rerio</i>) embryos. <i>Reproductive Toxicology</i> 33: 133-141.</p>	<p>The authors used a zebrafish embryo toxicity test (72 h exposure; specific malformation endpoints; proposed as teratogenicity screening test for human health to reduce reproductive toxicity studies) to study effects of coumarin and its derivative warfarin. Both caused teratogenic effects in zebrafish which demonstrates that the enzymes necessary for bioactivation are present in 72 h old zebrafish embryos (CYP2A6, CYP2C9 and others).</p>
<p>Incardona, J. P., T. L. Linbo and N. L. Scholz (2011). Cardiac toxicity of 5-ring polycyclic aromatic hydrocarbons is differentially dependent on the aryl hydrocarbon receptor 2 isoform during zebrafish development. <i>Toxicology and Applied Pharmacology</i> 257: 242-249.</p>	<p>The authors report on the role of the aryl hydrocarbon (AHR) / cytochrome P4501A (CYP1A) in the toxicity of different polycyclic aromatic hydrocarbons to very early life stages of zebrafish embryos (4-48 h post-fertilisation). PAHs induce CYP1A through AHR2 binding and cause cardiotoxic effects.</p>
<p>Kubota, A., J. J. Stegeman, B. R. Woodin, T. Iwanaga, R. Harano, R. E. Peterson, T. Hiraga and H. Teraoka (2011). Role of zebrafish cytochrome P450 CYP1C genes in the reduced mesencephalic vein blood flow caused by activation of AHR2. <i>Toxicology and Applied Pharmacology</i> 253(21): 244-252.</p>	<p>The study shows that TCDD induces CYP1Cs in zebrafish embryos (used up to 50 h post-fertilisation).</p>
<p><i>Reports on OECD validation of the zebrafish embryo toxicity test – Phase 1 & Phase 2</i> OECD (2011) OECD Series on Testing and Assessment No. 157: Validation report (Phase 1) for the zebrafish embryo toxicity test. Part 1 http://www.oecd.org/env/ehs/testing/48572244.pdf Part 2 http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=ENV/JM/MONO(2011)40&doclanguage=en</p> <p>OECD (2012b). Validation report (Phase 2) for the zebrafish embryo toxicity test. Series on Testing and Assessment No. 179: OECD, Paris, France. Summary http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282012%2925&doclanguage=en Annexes http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282012%2925/ANN&doclanguage=en</p>	<p>The development of TG 236 started in 2005 when the lead country for this TG, Germany, had submitted a first draft of the guideline with supporting background information to the OECD. OECD established an expert group to review the draft TG and the validation status of the test. Since reliable data on the transferability and reproducibility of the proposed protocol were lacking, it was decided to carry out a validation study to evaluate the performance of the test with 20 chemicals and as major deviations from the initially proposed protocol extend the exposure period from 48 to 96 h (beyond hatch) and increase the number of embryos from 10 to 20 / concentration and controls. The study was coordinated by EURL ECVAM and steered by a validation management group. Each chemical was tested in three independent runs in at least three laboratories. The results demonstrate that the test is transferable and reproducible. The protocol used in the validation study formed the basis for the final TG 236.</p>
<p>Lammer, E., G. J. Carr, K. Wendler, J. M. Rawlings, S. E. Belanger and T. Braunbeck (2009). Is the fish embryo toxicity test (FET) with</p>	<p>Lammer <i>et al.</i> compared LC50 values derived with fish embryo toxicity (FET) tests to LC50 values of acute fish toxicity (AFT) tests carried out with OECD TG 203 or other standard</p>

<p>the zebrafish (<i>Danio rerio</i>) a potential alternative for the fish acute toxicity test? <i>Comparative Biochemistry and Physiology</i> 149(2): 196-209.</p>	<p>tests. The FET database included 250 FET (mainly zebrafish, few fathead minnow, medaka and others) assays for 142 compounds. For 77 compounds AFT data from 2146 studies were available (restricted to main OECD species: rainbow trout, fathead minnow, zebrafish, medaka, bluegill). FET data had been retrieved from literature and AFT data from US EPA ECOTOX database and the ECETOC Aquatic toxicity database. Strict quality criteria had to be fulfilled for the FET: well-described protocol, exposure from 48-96 h, exposure via the water; peer-reviewed literature, in-house data from defined sources. Only AFT LC50 from 24-96 h exposure were used with preference for 96 h. Comparison of data pairs for 77 chemicals reveal a very good correlation of 0.90 and a slope of 1.07. They also compared AFT LC50s values of the five species and found correlations ranging from 0.89 to 0.96. The authors concluded that FET can provide equivalent information as AFT with juvenile and adult fish.</p>
<p>Knöbel, M., F. J. M. Busser, A. Rico-Rico, N. I. Kramer, J. L. M. Hermens, C. Hafner, K. Tanneberger, K. Schirmer and S. Scholz (2012). Predicting adult fish acute lethality with the zebrafish embryo: Relevance of test duration, endpoints, compound properties, and exposure concentration analysis. <i>Environmental Science and Technology</i> 46(17): 9690-9700.</p>	<p>This work was carried out within the framework of CELLsens project funded by CEFIC LRI and Defra (UK) [add reference]. It demonstrates the very good correlation of acute toxicity values (LC50) derived with zebrafish embryos and adult fish (fathead minnow [FHM] data from US EPA database) for 39 chemicals. For 12 chemicals values of other studies were used, whereas 27 chemicals were tested with zebrafish embryos following, in general, the first draft of the OECD FET guideline (OECD, 2006). Lethal and sublethal effects were recorded at 24h, 48h, and 120h (for at least one run/chemical). For 17 chemicals, exposure concentrations were measured. Comparison of nominal and measured LC50/EC50 zebrafish embryos and FHM revealed that measured concentrations improve the predictivity of the zebrafish embryo data for fish toxicity, whereas the prolongation of the exposure beyond hatch did not. Two chemicals were less toxic to embryos, the neurotoxic permethrin and ally alcohol requiring metabolic activation.</p>
<p>Belanger, S. E., J. M. Rawlings and G. J. Carr (2013). Use of fish embryo toxicity tests for the prediction of acute toxicity to chemicals. <i>Environmental Toxicology and Chemistry</i> 32(8): 1768-1783.</p>	<p>Belanger and colleagues increased the FET database initially compiled by Lammer <i>et al.</i> (2009) by results from recent FET studies, e.g. OECD validation study, Knöbel <i>et al.</i> (2012), unpublished data from in-house studies. The database used for the FET – AFT correlation consisted of 985 FET studies covering 229 compounds and 1531 AFT studies on 151 compounds. The range of log Kow was from -4.15 to 7.85, 15 different categories of functional use are evident with industrial organic compounds as largest category followed by pesticides, pharmaceuticals, surfactants and biocides. A total of 17 modes of action could be identified for the organic chemicals via ASTER. Compared to the Lammer <i>et al.</i> (2009) paper, the number of data pairs used for FET-AFT regression analysis doubled, i.e. from 77 to 151 chemicals. When including all data, correlation was $r = 0.90$ and slope of 1.027; when restricting exposure to 96 h for FET and AFT ($n = 72$), correlation was $r = 0.95$ and a slope</p>

	of 0.989. When comparing this relationship to all possible fish to fish relationships, the FET performs as a fish. The authors conclude that "In the FET is regarded as an alternative to the AFT, the FET will provide nearly equivalent predictions of hazard while improving overall animal welfare."
Henn, K. and T. Braunbeck (2011). Dechorionation as a tool to improve the fish embryo toxicity test (FET) with the zebrafish (<i>Danio rerio</i>). <i>Comparative Biochemistry and Physiology</i> 153C: 91-98.	Fish embryos are surrounded by the chorion, an acellular envelope until they hatch. It is not fully clarified whether the chorion presents a barrier and protects the embryo from exposure to chemicals. Henn and Braunbeck show that zebrafish embryos can be dechorionated at 24 h post-fertilisation with >90% survival rate and used in the zebrafish embryo toxicity test.
Mechanistic understanding of aquatic toxicity	
Verhaar, H. J. M., C. J. van Leeuwen and J. L. M. Hermens (1992). Classifying environmental pollutants. <i>Chemosphere</i> 25(4): 471-491.	In this paper a scheme is presented that makes it possible to classify a large number of organic pollutants into one of four classes: 1) inert chemicals, 2) less inert chemicals, 3) reactive chemicals and 4) specifically acting chemicals. For chemicals that are thus classified as belonging to one of these four classes it is possible to calculate either an expected effect concentrations, from a compound's octanol/water partition coefficient (Log Kow). For chemicals that cannot be classified as belonging to one of these four classes no prediction can be made. This approach can be implemented to estimate aquatic effect concentrations, which can be used to derive preliminary environmental quality objectives, or for the prioritisation of chemicals for subsequent testing. Moreover, these estimates could be of great value in risk and hazard assessment. This paper is especially focused on identifying the limits of applicability of Quantitative Structure-Activity Relationships for predicting aquatic toxicity.
Russom, C. L., S. P. Bradbury, S. J. Broderius, D. E. Hammermeister and R. A. Drummond (1997). Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (<i>Pimephales promelas</i>). <i>Environmental Toxicology and Chemistry</i> 16(5): 948-967.	In the field of aquatic toxicology, there has been an evolution of QSAR development and application from that of a chemical-class perspective to one that is more consistent with assumptions regarding modes of toxic action. The objective of this research was to develop procedures that relate modes of acute toxic action in the fathead minnow (<i>Pimephales promelas</i>) to chemical structures and properties. An empirically derived database for diverse chemical structures of acute toxicity and corresponding modes of toxic action was developed through joint toxic action studies, the establishment of toxicodynamic profiles, and behavioral and dose-response interpretation of 96 h LC50 tests. Using the results from these efforts, as well as principles in the toxicological literature, approximately 600 chemicals were classified as narcotics (three distinct groups), oxidative phosphorylation uncouplers, respiratory inhibitors, electrophiles/ proelectrophiles, acetylcholinesterase inhibitors, or central nervous system seizure agents. Using this data set, a computer-based expert system has been established whereby chemical structures are associated with likely modes of toxic action and, when available, corresponding QSARs.

<p>McKim, J. M., S. P. Bradbury and G. J. Niemi (1987). Fish acute toxicity syndrome and their use in the QSAR approach to hazard assessment. <i>Environmental Health Perspectives</i> 71: 171-186.</p>	<p>The authors describe the development of a "mode-of-action" database, and formulate a qualitative structure-activity relationship to assign the proper mode of action, and respective QSAR, to a given chemical structure. They define fish acute toxicity syndromes (FATS) as a characteristic set of whole-animal responses (clinical signs), based on various behavioural and physiological-biochemical measurements, and proposed that chemicals with a common mode of action will elicit a set of response associated with a specific FATS, thereby providing categories into which whole- fish response to acute toxicity can be grouped. Using behavioural parameters monitored in the fathead minnow during acute toxicity testing, FATS associated with acetylcholinesterase (AChE) inhibitors and narcotics could be reliably predicted. However, compounds classified as oxidative phosphorylation uncouplers or stimulants could not be resolved. Refinement of this approach by using respiratory-cardiovascular responses in the rainbow trout enabled FATS associated with AChE inhibitors, convulsants, narcotics, respiratory blockers, respiratory membrane irritants, and uncouplers to be correctly predicted.</p>
<p>Enoch, S. J., M. Hewitt, M. T. D. Cronin, S. Azam and J. C. Madden (2008). Classification of chemicals according to mechanism of aquatic toxicity: An evaluation of the implementation of the Verhaar scheme in Toxtree. <i>Chemosphere</i> 73(3): 243-248.</p>	<p>A number of mechanisms have been identified that can lead to (acute) aquatic toxicity, with the majority of industrial chemicals exerting their toxic influence via two non-covalent mechanisms: polar narcosis and non-polar narcosis. Verhaar et al. (1992) published a series of structural rules which aimed to classify compounds according to mechanism of action (see above); The importance of this classification scheme, and its utility for regulators and risk-assessors, has led it to be coded computationally within the Toxtree software from the European Chemicals Bureau website and within the OECD (Q)SAR Application Toolbox. The aim of this study was to assess the performance of the Verhaar classification scheme, using two important acute ecotoxicological databases, and to assess the strengths and weaknesses of this classification scheme. This study also highlights rule, and possible coding, errors that may lead to misclassifications. According to the authors, the Verhaar scheme in its current implementation misclassifies too many chemicals to be of significant practical use in terms of regulatory usage. The results suggest that significant additions are needed in the Verhaar scheme in order to improve the mechanistic assignments. According to the authors, such improvements should focus on a reordering of the rules in the decision tree, with the identification of electrophilic and specific noncovalent excess toxic mechanisms first (Verhaar classes 3 and 4, respectively), and the narcosis class (Verhaar class 1) should be the last class to be assigned. This is the opposite of the Verhaar approach and would enable a significant reduction in the numbers of chemicals currently assigned to Verhaar class 5. In addition, significant work on improving Verhaar classes 3 and 4 is required, involving the creation of more specific sub-classes, which would be extremely useful to aid</p>

	category formation and to enable mechanistic read-across predictions. Such additional subclasses for both Verhaar classes 3 and 4 should reflect the number of differing mechanisms of action that have been identified since the original Verhaar publication and would require rules for the identification of metabolically and abiotically activated chemicals. It is clear from the analysis presented that, should these changes be implemented, the classification results would be significantly improved.
ECETOC (2007)	<i>See below - Integrated testing strategies</i>
Ankley, G. T., R. S. Bennett, R. J. Erickson, D. J. Hoff, M. W. Hornung, R. D. Johnson, D. R. Mount, J. W. Nichols, C. L. Russom, P. K. Schmieder, J. A. Serrano, J. E. Tietge and D. L. Villeneuve (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. <i>Environmental Toxicology and Chemistry</i> 29(3): 730-741.	The authors discuss the use of the adverse outcome pathway (AOP) in ecotoxicology. An AOP is a conceptual construct that portrays existing knowledge concerning the linkage between a direct molecular initiating event and an adverse outcome at a biological level of organization relevant to risk assessment. AOP has been created to allow for an effective translation of this information into endpoints meaningful to ecological risk - effects on survival, development, and reproduction in individual organisms and, by extension, impacts on populations. This makes thus possible the use of mechanistic data to support chemical assessments, helping the ecological risk assessors to face the increasing demand to assess more chemicals, and to do so using fewer resources and experimental animals. The practical utility of AOPs for ecological risk assessment of chemicals is illustrated using five case examples. The examples demonstrate how the AOP concept can focus toxicity testing in terms of species and endpoint selection, enhance across-chemical extrapolation, and support prediction of mixture effects. The examples also show how AOPs facilitate use of molecular or biochemical endpoints for forecasting chemical impacts on individuals and populations. To conclude, the authors discuss how AOPs can help to guide research that supports chemical risk assessments and advocate for the incorporation of this approach into a broader systems biology framework.
OECD (2013c). Guidance Document on developing and assessing Adverse Outcome Pathways. Series on Testing and Assessment No. 184. ENV/JM/MONO(2013)6. OECD, Paris, France. Available from: http://www.oecd.org	This guidance document intends to provide the framework for consistent information gathering and organisation and to provide an introduction to the development and assessment of AOPs. It also provides initial assistance on how to undertake the assessment of an AOP in terms of its relevance and adequacy. A template has been included allowing authors to develop thorough AOPs and to improve consistency in AOPs developed by different risk assessor and stakeholders, as well as a glossary with definitions for AOP-specific terminology. The document also briefly outlines the potential use for regulatory purposes of AOP, but detailed guidance on how to use AOPs for integrated testing strategies and risk assessment will be developed in the future. A distinction is made between qualitative and quantitative AOP: a qualitative AOP is one where the key events have been identified but methods for assessing these events have not

	<p>been identified and/or assessed in sufficient detail to allow for identification of the applicability domains, threshold values and/or the response relationships to other key events. In contrast, a quantitative AOP is one where the methods for assessing the key events have been identified and sufficient data generated to identify the applicability domain, threshold values and/or the response relationships with other key events. An AOP may also be partial (i.e. one where not all key events are known); such an AOP may be useful in priority setting for further testing and development or may be used in hazard identification, as is currently performed with the OECD QSAR Toolbox. At present, physiologically-based pharmacokinetic (PBPK) modelling and toxicokinetics information on absorption, distribution, metabolism, and excretion (ADME) are not addressed in the AOP concept but will have to be addressed to develop a quantitative AOP required for a complete risk assessment.</p>
Status of non-standard methods	
QSARs	
<p>Pavan, M., A. Worth and T. I. Netzeva (2005a). Preliminary analysis of an aquatic toxicity dataset and assessment of QSAR models for narcosis. JRC Report EUR 21749 EN Available from: http://publications.jrc.ec.europa.eu/repository</p>	<p>In this work, analyses were performed to evaluate the possibility of using QSAR predictions for regulatory purposes. To this end, three literature-based narcosis QSAR models for acute fish toxicity to <i>Pimephales promelas</i> were analysed with respect to their ability to predict OECD Screening Information Data Set (SIDS) data for 177 High Production Volume (HPV) chemicals.</p> <p>The first two models are QSARs recommended by the EU Technical Guidance Document on chemical risk assessment for the polar and non-polar narcotic mechanisms of action. The third model was developed by ECB to represent the narcosis mechanism of action, including both non-polar and polar action. SIDS substances were classified according to expected Mode of Action (MOA) by using three different classification schemes and by applying a consensus scheme for the SIDS MOA. As a result, 75 chemicals were classified as non-polar narcotics (NPN) and 12 as polar narcotics (PN). Their acute toxicity to fish could be predicted confidently by the NPN and PN models. To predict the toxicities of the remaining 90 chemicals, which were classified as reactive, the use of MOA specific QSAR models was suggested. The three models were also assessed according to the extent to which they meet OECD principles for the validation of (Q)SARs for regulatory purposes, i.e. it should be associated with a defined endpoint, an unambiguous algorithm, a defined domain of applicability, appropriate measures of internal performance (as represented by goodness-of-fit and robustness), predictivity (as determined by external validation), and a mechanistic interpretation (wherever possible). Model predictions for the 57 substances ranged from 0.1–10 times the actual acute toxicity values, when the model was applied only to substances in</p>

	<p>the model's applicability domain. The range became much larger when the applicability domain constraints were ignored. Those results emphasise the importance of defining the model applicability domain in order to identify reliable predictions for the regulatory assessment of chemicals.</p>
<p>Pavan, M., A. Worth and T. I. Netzeva (2005b). Comparative assessment of QSAR models for aquatic toxicity. JRC Scientific and Technical Report EUR 21750 EN Available from: http://publications.jrc.ec.europa.eu/repository</p>	<p>This report describes a comparative assessment of QSAR models for aquatic toxicity, in order to evaluate the possibility of using QSAR predictions for regulatory purposes. Six literature-based QSAR models for acute fish toxicity were analysed with respect to their ability to predict OECD Screening Information Data Set (SIDS) data for 177 High Production Volume (HPV) chemicals.</p> <p>The first three models are the ones quoted above (Pavan et al., 2005a); the fourth model is more general than the previous ones since by including an electrophilicity descriptor it is supposed to describe potentially bioreactive (electrophilic) chemicals; the fifth model is based on hydrophobic and polar atom-type electrotopological state (E-state) indices; and the sixth model is a commercially available neural network software program, developed by the TerraBase Inc., for the computation of acute (96hr) median lethal concentrations (LC50) of organic substances. The SIDS substances were classified according to expected Mode of Action according to a consensus classification scheme based on three different schemes. The six models were also assessed according to the extent to which they meet OECD principles for the validation of (Q)SARs for regulatory purposes (see above). For each model, a comparison between predictions and experimental fish toxicity was performed by calculating the number of chemicals with predicted effect concentrations within factors of 10, 100 and 1000 of the corresponding SIDS test data. For each model the ratio was calculated first by using the entire SIDS data set and then by using only the chemicals falling with the model applicability domain. The results show that when the model domain was taken into account, the ratio was always near one and in the range from 0.1 to 10. The range became much larger when the applicability domain constraints were ignored. The results of this study support the view that the regulatory application of a QSAR model should be based on a suitable definition of the model applicability domain in order to identify reliable predictions.</p>
<p>Netzeva, T. I., M. Pavan and A. Worth (2007). Review of data sources, QSARs, and integrated Testing Strategies for aquatic toxicity. European Commission Joint Research Centre Scientific and Technical Report, EUR 229443EN, Ispra, Italy. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository</p>	<p>This review collects information on sources of aquatic toxicity data and computational tools for estimation of chemical toxicity to aquatic organisms, such as expert systems and quantitative structure-activity relationship (QSAR) models. The review also captures current thinking of what constitutes an integrated testing strategy (ITS) for this endpoint and put an emphasis on the usefulness of the models for the regulatory assessment of chemicals, particularly for the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) legislation. Effects on organisms from three trophic levels (fish, Daphnia and</p>

	<p>algae) are considered and data sources such as databases and papers publishing experimental data are identified. Models for narcosis, general (global) models as well as models for specific chemical classes and mechanisms of action are summarised. Where possible, models were included in a form allowing reproduction without consultation with the original paper. This review builds on work carried out in the framework of the REACH Implementation Projects, and was prepared as a contribution to the EU funded Integrated Project, OSIRIS.</p>
<p>Netzeva, T. I., M. Pavan and A. P. Worth (2008). Review of (Quantitative) Structure–Activity Relationships for Acute Aquatic Toxicity. <i>QSAR & Combinatorial Science</i> 27(1): 77-90.</p>	<p>This paper reviews different approaches described in the literature for estimating the aquatic toxicity of chemical substances. It is based on an extended review performed by the European Chemicals Bureau of the European Commission's Joint Research Centre in support of the development of technical guidance for the implementation of the REACH legislation. The paper is organised by approach for (Q)SAR development and includes a review of: (i) (Q)SARs for acute aquatic toxicity by chemical class, (ii) (Q)SARs for acute aquatic toxicity by mode of action, (iii) review of statistically derived (Q)SARs, (iv) structural alerts for excess aquatic toxicity and (v) expert systems that combine structural rules and multiple (Q)SAR models to predict aquatic toxicological endpoints. Effects on organisms from three trophic levels (fish, Daphnia and algae) were considered, and both traditional data sources such as databases and literature were identified. Models for narcosis, general (global) models as well as models for specific chemical classes and mechanisms of action were summarised. This review also captures current thinking of what constitutes an integrated testing strategy (ITS) for this endpoint, and put an emphasis on the usefulness of the models and for the regulatory assessment of chemicals. Directions and recommendations for further research are also provided.</p>
<p>Brooke, D. and M. Crookes (2006a). Validation of (quantitative) structure-activity relationships for toxicological endpoints of regulatory importance: (Q)SARs for acute toxicity to fish - Part A. European Commission - Joint Research Centre - Institute for Health and Consumer Protection: 393 p. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_BRE_partA.pdf</p>	<p>The objective of this project was to identify a short-list of promising QSARs for acute toxicity to fish, and to provide sufficient information on each short-listed QSAR to enable its scientific validity to be established.</p> <p>This report reviewed and summarized 578 QSARs for acute fish toxicity from 81 sources, and provides recommendations for a short-list of 7 promising QSARs that should be further validated, as well as for the aspects that should be considered during the further validation of QSARs and the application of QSARs. The results of this validation exercise are given in a separate report (Part B). All of these QSARs consider the 96h LC50 (amongst other endpoints).</p> <p>Based on this work, recommendations for future approaches for validation of QSARs for fish toxicity are also provided. This include recommendations for the factors to be considered when validating QSARs and consideration of the use of <i>in vitro</i> fish cell line tests for the development of high quality QSARs and/or identification of the mode of toxic action of new</p>

<p>Brooke, D. and M. Crookes (2006b). Validation of (quantitative) structure-activity relationships for toxicological endpoints of regulatory importance: (Q)SARs for acute toxicity to fish - Part B. European Commission - Joint Research Centre - Institute for Health and Consumer Protection: 138 p. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_BRE_partB.pdf</p>	<p>unknown chemical</p> <p>The following further validation work has been carried out (where possible) on the 7 short-listed QSARs in Part A (above):</p> <ul style="list-style-type: none"> - Verification of the algorithm and associated statistics given in the paper. - Cross-validation of the QSAR. - Validation using an independent test set. - Exploration of the domain of applicability of the QSAR. <p>This work has shown that the confidence in the validity of QSARs for acute fish toxicity can be increased by carrying out relatively simple analysis of the QSARs. The score of all of the QSARs against the OECD principles for QSAR validity (given in Part A of this report) have been substantially increased.</p>
<p>EFSA (2013). Scientific Opinion - Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Panel on Plant Protection Products and their Residues (PPR). EFSA Journal 11(7): 3290. http://www.efsa.europa.eu/en/efsajournal/doc/3290.pdf</p>	<p>Regarding the use of QSARs, this guidance states the following (executive summary):</p> <p><i>Guidance, largely following the European Chemicals Agency (ECHA) recommendations (ECHA, 2008), is provided on the use of non-testing methods in PPP RA, such as (Quantitative) Structure– Activity Relationship ((Q)SAR) models, expert test systems and analogue read-across as tools for deriving intrinsic properties of substances. Non-testing methods may be used to estimate endpoints for metabolites without the toxophore and for impurities. In addition, (Q)SARs might, together with available test data, be used to rank species for identifying the most likely sensitive taxonomic group to focus experimental testing (EFSA PPR Panel, 2012a). For a detailed description of non-testing methods see section 10.1.</i></p> <p><i>Only suitable models (e.g. covering the right domain) with a high predictive reliability should be used (see section 10.1.2). This should, among others, be reflected in the level of statistical significance required for estimates from (Q)SAR models. Validation parameters should ideally indicate good fits (e.g. $Q_2 > 0.7$, concordance correlation coefficient (CCC) > 0.85). Estimates of toxicity should, where possible, be assisted by confidence intervals around the prediction. In case the standard deviation exceeds the predicted value itself, such values should not be accepted. Generally, the worst-case endpoint from several modelling approaches should be used.</i></p> <p><i>Estimates should be confirmed by using weight-of-evidence approaches where all available information is taken into account. This could include a combination of the different (Q)SAR model predictions combined with read-across and other available information like non-standard test data and TK/TD information from mammals.</i></p> <p><i>To date, most experience is gained with (Q)SAR models that predict acute toxicity. It is noted that fewer valid (Q)SAR models are currently available for deriving chronic toxicity data.</i></p>

	<p>A decision scheme for use of non-testing systems is presented below (see section 10.1). Decision scheme not included here.</p>
<p>Pavan, M., T. I. Netzeva and A. Worth (2006). Validation of a QSAR model for acute toxicity. SAR and QSAR in Environmental Research 17: 147-171.</p>	<p>In this work, a quantitative structure – activity relationship (QSAR) model has been developed for predicting acute toxicity to the fathead minnow (<i>Pimephales promelas</i>) in order to demonstrate how statistical validation and domain definition are both required to establish model validity and to provide reliable predictions. A dataset of 408 heterogeneous chemicals was modeled by a diverse set of theoretical molecular descriptors by using multivariate linear regression (MLR) and Genetic Algorithm – Variable Subset Selection (GA-VSS). This QSAR model was developed to generate reliable predictions of toxicity for organic chemicals not yet tested, so particular emphasis was given to statistical validity and applicability domain. External validation was performed by using OECD Screening Information Data Set (SIDS) data for 177 High Production Volume (HPV) chemicals, and a good predictivity was obtained ($Q^2_{ext} = 72.1$). The model was evaluated according to the OECD principles for QSAR validation, and compliance with all five principles was established. According to the author, the model could therefore be useful for the regulatory assessment of chemicals, and could be used to fill data gaps within its chemical domain or contribute to the prioritization of chemicals for aquatic toxicity testing.</p>
<p>Cassani, S., S. Kovarich, E. Papa, P. P. Roy, L. van der Wal and P. Gramatica (2013). Daphnia and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity-activity modelling. Journal of Hazardous Materials 258-259: 50-60.</p>	<p>In the present study, QSAR models for the prediction of acute synthetic triazoles and benzo-triazoles ((B)TAZs) toxicity in <i>Daphnia magna</i> and <i>Onchorhynchus mykiss</i> have been developed according to the principles for the validation of QSARs and their acceptability for regulatory purposes, proposed by the Organisation for Economic Co-operation and Development(OECD). They have been applied to predict acute toxicity for over 300 (B)TAZs without experimental data. Additionally, on the basis of the high correlation coefficient found among the experimental values of pEC(LC)50 available for <i>Daphnia magna</i> and <i>Onchorhynchus mykiss</i> ($r=0.93$), a model based on Quantitative Activity-Activity Relationships (QAAR) has been developed, which allows for interspecies extrapolation from daphnids to fish. On 40 substances, only one response outlier was found (the most simple and least toxic compound, 1,2,4-triazole), and this QAAR model overestimate its toxicity. According to the authors, predictions generated by the here proposed QSAR and QAAR models for all the studied chemicals can be used by regulators to support the use of Weight-of-Evidence and non testing-based approaches. The authors also highlight the importance of QSAR/QAAR, especially when dealing with specific chemical classes like (B)TAZs, for screening and prioritization of pollutants under REACH.</p>
<p>Jacobs, M. N. (2004). In silico tools to aid risk assessment of endocrine disrupting chemicals. Toxicology 205(1–2): 43-53.</p>	<p>This paper reviews <i>in silico</i> tools used in risk assessment of endocrine disrupting chemicals, including the use of: (1) nuclear receptor (NR) crystal structures and homology models to</p>

	<p>examine potential modes of ligand binding by different representative compounds; (2) multivariate principal component analyses (PCA) techniques to select best predicted cell lines for endocrine disrupting chemicals (EDC) risk assessment purposes; (3) NR quantitative structure-activity relationships (QSARs) that can be constructed from varied biological data sources. The cytosolic and NR examples discussed in this work include the Ah receptor, (AhR), the human oestrogen receptor α (hERα) and the human pregnane X receptor (PXR). The varied biological data sets can be compared to give a more integrated dimension to receptor cross talk mechanisms, with further support from molecular modelling studies. According to the author, <i>in silico</i> or computational tools could be used more effectively in endocrine disruptor risk assessment for prescreening potential endocrine disruptors, improving experimental <i>in vitro</i> screening assay design and facilitating more thorough data analyses.</p>
<p>de Haas, E. M., T. Eikelboom and T. Bouwman (2011). Internal and external validation of the long-term QSARs for neutral organics to fish from ECOSAR™. SAR and QSAR in Environmental Research 22(5-6): 545-559.</p>	<p>This study concentrates on the external validation of an existing Quantitative Structure–Activity Relationship (QSAR) model widely used for long-term aquatic toxicity to fish. The predictivity of the model was evaluated in order to increase its reliability. The authors assessed whether the model met all of the OECD principles. The model was adapted to become more robust, and predictions were made with an external validation set collected from several databases. For the internal validation of the QSAR, the r^2, Q^2_{LOO} and Q^2_{LMO} were used as validation criteria, and for the external validation r^2, Q^2_{ext}, h and the validation ratio were used. A few substances were classified as outliers and therefore the applicability domain of the QSAR had to be adjusted. The QSAR passed all validation criteria and met all the OECD principles for QSAR validation. According to the authors, the long-term toxicity QSAR for fish can be applied with high certainty of a correct prediction within the limits of the inherent uncertainty of the model in cases where the substance falls within the applicability domain.</p>
<p>Claeys, L., Iaccino, F., Janssen, C.R., Van Sprang, P., Verdonck, F. (2013) Development and validation of a quantitative structure-activity relationship for chronic narcosis to fish. Environmental Toxicology and Chemistry 32(10): 2217-25.</p>	<p>The authors describe the development of a QSAR for prediction of chronic fish toxicity of MOA1 and MOA2 chemicals using NOEC values derived from mortality endpoint. The MOA1 model is based on a training set of 49 chemicals (114 experimental data points) and a test set of 20 chemicals: $\log \text{NOEC}(\text{mmol/L}) = -0.9402 \log \text{Kow} + 0.8911$. Internal validation provided the following values: $r^2 = 0.76$, $r^2_{adj} = 0.76$, $Q^2_{LOO} = 0.75$. External validation reveals $r^2_{ext} = 0.72$ confirming the good predictivity of the model. The MOA2 model is based on a training set of 10 chemicals (42 experimental data points) and a test set of 3 chemicals: $\log \text{NOEC}(\text{mmol/L}) = -1.0117 - 0.61147 \log \text{Kow} + 0.04177$. Internal validation provided the following values: $r^2 = 0.80$, $r^2_{adj} = 0.79$, $Q^2_{LOO} = 0.78$ and external validation gives an $r^2_{ext} = 0.65$.</p>

<p>Austin, T.J. and C.V. Eadsforth (2014). Development of a chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics. SAR QSAR Environmental Research 25(2): 147-60.</p>	<p>The authors developed a QSAR using only sub-lethal no-observed-effect concentration (NOEC) end-point data according to best practice QSAR development. Only the lowest NOEC value was taken for each compound, in line with the conservative approach taken by the ECHA.</p> <p>Using a training set of 19 chemicals a model was created: $\log \text{NOEC}(\text{mmol/L}) = 0.711 - 0.914 \log \text{Kow}$. Internal validation resulted in the following values: $r^2 = 0.91$, $r^2_{\text{adj}} = 0.9$, $Q^2_{\text{LOO}} = 0.88$. For external validation a set of 10 chemicals was used and the $r^2_{\text{ext}} = 0.89$ confirms the good predictivity of the model. The authors conclude that the model developed meets the OECD principles, has strong internal and external validation statistics, and can reliably predict sub-lethal chronic NOEC values for fish within its defined applicability domain, i.e. chemicals which act via non-polar narcosis and have a logKow between 0.46 and 5.30.</p>
<p>Schultz, T. W. (2010). Chapter 14 Adverse Outcome Pathways: A Way of Linking Chemical Structure to In Vivo Toxicological Hazards. In Silico Toxicology: Principles and Applications, The Royal Society of Chemistry: 346-371.</p>	<p>The concept of Adverse Outcome Pathway (AOP) and its potential to link chemical structure of toxicants to <i>in vivo</i> toxicological hazards is described. A definition of a chemical category in hazards assessment (a group of chemicals whose physico-chemical properties, human health and/or environmental toxicological properties and/or environmental fate properties are likely to be similar or follow a regular pattern) is given and their importance to fill data gap, whether by read-across or quantitative structure-activity relationship (QSAR) modelling explained. The authors also highlight the importance of the mode of action within a category of chemical, based on the fact that the compounds are grouped on the hypothesis that the properties of the chemicals in the category will show coherent trends in their toxicological effects. However, current knowledge of toxicological categories and category formation is limited, in part due to the depth and breadth of available database needed to support category formation and the complexity of the hazard endpoint being evaluated. The authors also present the AOP concept and give some examples of detailed AOPs, such as skin sensitisation, acetylcholine esterase inhibition, or receptor binding pathways for phenolic oestrogen mimics. They discuss advantages of the AOP approach and to conclude with the basic elements in developing a pathway.</p>
<p>Tremolada, P., A. Finizio, S. Villa, C. Gaggi and M. Vighi (2004). Quantitative inter-specific chemical activity relationships of pesticides in the aquatic environment. Aquatic Toxicology 67: 87-103.</p>	<p>This work aims at developing quantitative inter-specific chemical activity relationship for aquatic organisms to verify if such an approach could be utilised for estimating toxicological data when no other information is available. Inter-specific toxicity relationships on fish, <i>Daphnia</i>, and algae were performed for pesticides considering more than 600 compounds. Good correlation were found between several fishes species and were improved by excluding highly specific compounds such as organophosphorous insecticides.</p>
<p>de Roode, D., C. Hoekzema, S. de Vries-Buitenweg, B. van de Waart and J. van der Hoeven (2006). QSARs in ecotoxicological</p>	<p>In this paper, four QSARs were evaluated to predict toxicity for 170 compounds from a broad chemical class, using them as a black-box. Predictions were obtained for 122</p>

<p>risk assessment. <i>Regulatory Toxicology and Pharmacology</i> 45(1): 24-35.</p>	<p>compounds, indicating an important drawback of QSARs, i.e., for 28% of the compounds QSARs cannot be used at all. ECOSAR, Topkat, and QSARs for non-polar and polar narcosis generated predictions for 120, 39, 24, and 11 compounds, respectively. Correlations between experimental and predicted effect concentrations were significant for Topkat and the QSAR for polar narcosis, but generally poor for ECOSAR and the QSAR for non-polar narcosis. When predicted effect concentrations for fish were allowed to deviate from experimental values by a factor of 5, correct predictions were generated for 77%, 54%, 68%, and 91% of the compounds using ECOSAR, Topkat, and the QSARs for non-polar and polar narcosis, respectively. It was impossible to indicate specific chemical classes for which a QSAR should be used or not. Those results show that currently available QSARs cannot be used as a black-box.</p>
<p>Yan, D., X. Jiang, G. Yu, Z. Zhao, Y. Bian and F. Wang (2006). Quantitative structure–toxicity relationships of organophosphorous pesticides to fish (<i>Cyprinus carpio</i>). <i>Chemosphere</i> 63(5): 744-750.</p>	<p>This study was conducted to determine the relationships between 1381 chemical and structural parameters of 43 organophosphorus pesticides (OPs) and their toxicity to fish, <i>Cyprinus carpio</i>. By multivariate linear regression and intervariable regression analyses, various equations have been derived to calculate the lethal toxicity value, LC50, for 43 OPs found in fish with different levels of toxicity. Results show that for all selected OPs, especially those of low toxic OPs ($LC_{50} < 2.5$ mM), one equation could account for 86.2% of the variability of the toxic effect. The steric and electronic characteristics and the hydrophobicity of OPs, in particular, are among the most important parameters determining the toxicity of OPs to fish. For the OPs with high toxicity, different structural parameters were introduced into two other equations. These results suggest that chemical and structural parameters could be useful in modelling chemical reactivity within homologous series of OP compounds and elucidating possible mechanisms associated with different levels of toxicity to fish.</p>
<p>Yuan, H., Y.-Y. Wang and Y.-Y. Cheng (2007). Mode of action-based local QSAR modeling for the prediction of acute toxicity in the fathead minnow. <i>Journal of Molecular Graphics and Modelling</i> 26(1): 327-335.</p>	<p>In this paper, the acute toxicity, 96 h LC50 (median lethal concentration) for the fathead minnow served as the toxicity endpoint of interest, and the mode of action (MOA) was employed as a criterion to compartmentalize the chemical domains. MOA-based local QSAR models were built by partial least squares (PLS) regression for each subset with single mode of action such as Narcosis I, Narcosis II or Reactive, and global model was also developed for the combined data set containing several subsets above. By comparing the performances of these two types of models, the local models were superior to the global model in that the relative standard error (R.S.E.) of the former was much lower for both the training set and the test set of any subset. In addition, the influence of the reliability of MOA determination on the performance of local model was also investigated and the statistical results for subsets with MOAs at A and B confidence level were better than those at C and D confidence level.</p>

	Therefore, the MOA-based local QSAR models are promising to improve the accuracy of toxicity prediction as long as the assessment of MOA is of high reliability.
Colombo, A., E. Benfenati, M. Karelson and U. Maran (2008). The proposal of architecture for chemical splitting to optimize QSAR models for aquatic toxicity. <i>Chemosphere</i> 72(5): 772-780.	In QSAR analysis, compounds have often been divided into distinct groups according to their mode of action or chemical class. In the current study, theoretical molecular descriptors were used to divide 568 organic substances into subsets with toxicity measured for the 96 h lethal median concentration for the fathead minnow (<i>Pimephales promelas</i>). Simple constitutional descriptors such as the number of aliphatic and aromatic rings and a quantum chemical descriptor, maximum bond order of a carbon atom divide compounds into nine subsets. For each subset of compounds the automatic forward selection of descriptors was applied to construct QSAR models. Significant correlations were achieved for each subset of chemicals and all models were validated with the leave-one-out internal validation procedure ($r^2_{cv} \approx 0.80$). The results encourage considering this alternative way for the prediction of toxicity using QSAR subset models without direct reference to the mechanism of toxic action or the traditional chemical classification.
Reuschenbach, P., M. Silvani, M. Dammann, D. Warnecke and T. Knacker (2008). ECOSAR model performance with a large test set of industrial chemicals. <i>Chemosphere</i> 71(10): 1986-1995.	This work aimed at evaluating the widely used ECOSAR computer programme for QSAR prediction of chemical toxicity towards aquatic organisms. This was done by using large data sets of industrial chemicals with varying molecular structures. Experimentally derived toxicity data covering acute effects on fish, <i>Daphnia</i> and green algae growth inhibition of more than 1000 randomly selected substances were compared to the prediction results of the ECOSAR programme in order (1) to assess the capability of ECOSAR to correctly classify the chemicals for aquatic toxicity according to rules of EU regulation and (2) to determine the number of correct predictions within tolerance factors from 2 to 1000. Regarding ecotoxicity classification, 65% (fish), 52% (<i>Daphnia</i>) and 49% (algae) of the substances were correctly predicted into the classes “not harmful”, “harmful”, “toxic” and “very toxic”. At all trophic levels about 20% of the chemicals were underestimated in their toxicity. The class of “not harmful” substances (experimental LC/EC50 > 100 mg l ⁻¹) represents nearly half of the whole data set. The percentages for correct predictions of toxic effects on fish, <i>Daphnia</i> and algae growth inhibition were 69%, 64% and 60%, respectively, when a tolerance factor of 10 was allowed. Focussing on those experimental results which were verified by analytically measured concentrations, the predictability for <i>Daphnia</i> and algae toxicity was improved by approximately three percentage points, whereas for fish no improvement was determined. The calculated correlation coefficients demonstrated poor correlation when the complete data set was taken, but showed good results for some of the ECOSAR chemical classes. The authors also discuss the results in the context of literature data on the performance of ECOSAR and other QSAR models.

<p>Jager, T. and S. L. M. Kooijman (2009). A biology-based approach for quantitative structure-activity relationships (QSARs) in ecotoxicity. <i>Ecotoxicology</i> 18(2): 187-196.</p>	<p>The standard approach in developing QSARs for toxicity is to collect toxicity values for one species for a group of chemicals, and attempt to find one or a few molecular descriptors that, in some form of regression, provide an adequate description. Progress in this field has been limited to developing equations for new species, new groups of toxicants and using other descriptors. However, according to the authors, the use of descriptive summary statistics for toxicity, such as the 4-day LC50 for fish, introduces bias and ignores valuable kinetic information in the data. To extract all relevant information from toxicity test results requires biology-based methods such as DEBtox. Those methods use all of the toxicity data in time to derive time-independent and unbiased parameter estimates. Additionally, those parameters are expected to co-vary in specific ways, which offers unique opportunities for the development of predictive QSARs. In this paper, the authors explore the potential of biology-based modelling in QSARs development, by applying the hazard model from DEBtox to analyse survival data for fathead minnows (<i>Pimephales promelas</i>), and they demonstrate how these methods can lead to a different approach towards QSARs. As a conclusion, the results show that different modes of action resulted in different patterns in the parameter estimates, and therefore, that the toxicity data by themselves reveal insight into the actual mechanism of toxic action.</p>
<p>Sanderson, H. and M. Thomsen (2009). Comparative analysis of pharmaceuticals versus industrial chemicals acute aquatic toxicity classification according to the United Nations classification system for chemicals. Assessment of the (Q)SAR predictability of pharmaceuticals acute aquatic toxicity and their predominant acute toxic mode-of-action. <i>Toxicology Letters</i> 187(2): 84-93.</p>	<p>The aim of this paper was to compile a comprehensive database based on OECD's standardised measured ecotoxicological data and to evaluate if there is generally cause of greater concern with regard to pharmaceutical aquatic toxicological profiles compared to industrial chemicals. Comparisons were based upon aquatic ecotoxicity classification under the United Nations Global Harmonized System for classification and labeling of chemicals (GHS). The authors statistically explored whether the predominant mode-of-action (MOA) for pharmaceuticals is narcosis, and found 275 pharmaceuticals with 569 acute aquatic effect data and 23 pharmaceuticals with chronic data. Pharmaceuticals were found to be more frequent than industrial chemicals in GHS category III. Acute toxicity was predictable (>92%) using a generic (Q)SAR ((Quantitative) Structure Activity Relationship) suggesting a narcotic MOA. Analysis of model prediction error suggests that 68% of the pharmaceuticals have a non-specific MOA. Additionally, the acute-to-chronic ratio (ACR) for 70% of the analysed pharmaceuticals was below 25 further suggesting a non-specific MOA. Sub-lethal receptor-mediated effects may however have a more specific MOA.</p>
<p>Dom, N., D. Knapen and R. Blust (2012). Assessment of aquatic experimental <i>versus</i> predicted and extrapolated chronic toxicity data of four structural analogues. <i>Chemosphere</i> 86(1): 56-64.</p>	<p>The present study was developed to assess the chronic toxicity predictions and extrapolations for a set of chlorinated anilines (aniline (AN), 4-chloroaniline (CA), 3,5-dichloroaniline (DCA) and 2,3,4-trichloroaniline (TCA)). <i>Daphnia magna</i> 21 days chronic experimental data was compared to the chronic toxicity predictions made by the US EPA ECOSAR QSAR</p>

	<p>tools and to acute-to-chronic extrapolations. Additionally, Species Sensitivity Distributions (SSDs) were constructed to assess the chronic toxicity variability among different species and to investigate the acute <i>versus</i> chronic toxicity in a multi-species context. Although chlorinated anilines are structural analogues with a designated polar narcotic mode of action, rather large interchemical and interspecies differences in toxicity were observed. Compared to the other three test compounds, TCA exposure had a significantly larger impact on growth and reproduction of <i>D. magna</i>. Furthermore, this study illustrated that QSARs or a fixed ACR are not able to account for these interchemical and interspecies differences. Consequently, ECOSAR was found to be inadequate to predict the chronic toxicity of the anilines and the use of a fixed ACR (of 10) led to under of certain species. The experimental ACRs determined in <i>D. magna</i> were substantially different among the four aromatic amines (ACR of 32 for AN, 16.9 for CA, 5.7 for DCA and 60.8 for TCA). Furthermore, the SSDs illustrated that <i>Danio rerio</i> was rather insensitive to AN in comparison to another fish species, <i>Pimephales promelas</i>. The authors therefore suggest that available toxicity data should be used in an integrative multi-species way, rather than using individual-based toxicity extrapolations. In this way, a relevant overview of the differences in species sensitivity is given, which in turn can serve as the basis for acute to chronic extrapolations.</p>
<p>McCarty, L. S. (2012). Model validation in aquatic toxicity testing: Implications for regulatory practice. <i>Regulatory Toxicology and Pharmacology</i> 63(3): 353-362.</p>	<p>In this study, the authors wanted to check the validity of models and assumptions used in toxicity testing. In this aims, they have done a quality control evaluation of the acute toxicity testing protocol using the US. EPA fathead minnow database, focusing around three key assumptions that ensure results represent valid toxicological metrics: 1) it must be possible to estimate steady-state LC50s; 2) LC50s should occur at equivalent exposure durations; 3) all substantive toxicity modifying factors should be adequately controlled. About 8% of the tests failed the first assumption and are invalid and unusable. Examination of remaining data indicated variance from unquantified effects of toxicity modifying factors remained in LC₅₀s, thereby failing assumption three. Such flaws in toxicity data generated via recommended LC50 testing protocols mean that resulting data do not represent consistent, comparable measures of relative toxicity. Then, according to the authors, current regulations employing LC50 testing data are acceptable due to the use of semiquantitative, policy-driven development guidance that considers such data uncertainty, but quantitative applications such as QSARs, mixture toxicity, and regulatory chemical grouping can be compromised. These validation failures justify a formal quality control review of the LC50 toxicity testing protocol. Interim improvements in the design, execution, interpretation, and regulatory applications of LC50 and related protocols using exposure-based dose surrogates are warranted.</p>

Toxicokinetic models	
<p>Stadnicka, J., K. Schirmer and R. Ashauer (2012). Predicting concentrations of organic chemicals in fish by using toxicokinetic models. <i>Environmental Science & Technology</i> 46: 3273–3280.</p>	<p>In aquatic toxicology, quantification of chemical toxicity is usually based on nominal or measured concentrations (here named external concentration) of the chemical in the water. However, chemical concentrations in the fish (here named internal concentration) have been suggested to be a more suitable parameter. To better understand the relationship between the external and internal concentrations of chemicals in fish, and to quantify internal concentrations, the authors compared three toxicokinetic (TK) models with each other and with literature data of measured concentrations of 39 chemicals. They used two one compartment models, a physiologically based toxicokinetic (PBTK) model, improved for the treatment of lipids, to predict concentrations of organic chemicals in two fish species: rainbow trout (<i>Oncorhynchus mykiss</i>) and fathead minnow (<i>Pimephales promelas</i>). The three models predicted the measured internal concentrations in fish within 1 order of magnitude for at least 68% of the chemicals. The PBTK model outperformed the one-compartment models with respect to simulating chemical concentrations in the whole body (at least 88% of internal concentrations were predicted within 1 order of magnitude using the PBTK model). All the models can be used to predict concentrations in different fish species without additional experiments. However, further development of TK models is required for polar, ionisable, and easily biotransformed compounds.</p>
<p>Stadnicka-Michalak, J., Tanneberger, K., Schirmer, K., et al., 2014. Measured and modeled toxicokinetics in cultured fish cells and application to in vitro-in vivo toxicity extrapolation. <i>PLoS One</i>. 9, e92303.</p>	<p>Effect concentrations in the toxicity assessment of chemicals with fish and fish cells are generally based on external exposure concentrations. External concentrations as dose metrics, may, however, hamper interpretation and extrapolation of toxicological effects because it is the internal concentration that gives rise to the biological effective dose. There is a need to understand the relationship between the external and internal concentrations of chemicals and therefore the authors investigated the following: (i) the time-course of the concentration of chemicals with a wide range of physicochemical properties in the compartments of an in vitro test system, (ii) derive a predictive model for toxicokinetics in the in vitro test system, (iii) test the hypothesis that internal effect concentrations in fish (in vivo) and fish cell lines (in vitro) correlate, and (iv) develop a quantitative in vitro to in vivo toxicity extrapolation method for fish acute toxicity. To achieve these goals, time-dependent amounts of organic chemicals were measured in medium, cells (RTgill-W1) and the plastic of exposure wells. Then, the relation between uptake, elimination rate constants, and log Kow was investigated for cells in order to develop a toxicokinetic model. This model was used to predict internal effect concentrations in cells, which were compared with internal effect concentrations in fish gills predicted by a Physiologically Based Toxicokinetic model. The model could predict concentrations of non-volatile organic chemicals with log Kow</p>

	between 0.5 and 7 in cells. The correlation of the log ratio of internal effect concentrations in fish gills and the fish gill cell line with the log Kow was significant ($r > 0.85$, $p = 0.0008$, Ftest). This ratio can be predicted from the log Kow of the chemical (77% of variance explained), comprising a promising model to predict lethal effects on fish based on in vitro data.
<i>Reduction - Step-down Approach – basis for the above described threshold approach</i>	
Hutchinson, T. H., S. Barret, M. Buzby, D. Constable, A. Hartmann, A. Hayes, D. Huggett, R. Länge, A. D. Lillicrap, J. O. Straub and R. S. Thompson (2003). A strategy to reduce the numbers of fish used in acute ecotoxicity testing of pharmaceuticals. <i>Environmental Toxicology and Chemistry</i> 22: 3031-3036.	At the time of the publication of this paper, the draft European regulation on environmental safety assessment of human pharmaceuticals was discussed. The authors proposed that testing can be moved from fish LC50 testing (typically using 42 fish/active pharmaceutical ingredient [API]) to acute threshold tests using fewer fish (typically 10 fish/API). Base-set ecotoxicity data from regulatory studies of 91 API were collated and found that for 73 of the 91 APIs, the algal EC50 and daphnid EC50 values were lower than or equal to the fish LC50 data. Thus, for approximately 80% of these APIs, algal and daphnid acute EC50 data could have been used in the absence of fish LC50 data to derive PNECwater values. For the other 18 APIs, use of an acute threshold test with a step-down factor of 3.2 is predicted to give comparable PNECwater outcomes. Based on this so-called fish acute threshold (step-down) approach the total number of fish used could be reduced from 3,822 to 1,025 (73%) with no loss of data for PNECwater estimates.
Jeram, S., J. M. R. Sintes, M. Halder, J. B. Fentanes, B. Sokull-Klüttgen and T. H. Hutchinson (2005). A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. <i>Regulatory Toxicology and Pharmacology</i> 42(2): 218-224.	The authors applied the above described threshold/step-down approach to industrial chemicals from the New Chemicals Database (1439 data sets). The findings of Hutchinson et al. (2003) were confirmed and only for 15-20% chemicals fish was the most sensitive test species. By applying the threshold (step-down) approach a reduction of at least 50%-70% would be feasible. In contrast to Hutchinson et al. (2003), the authors propose to use seven fish/concentration and control, which corresponds to the minimum number of fish required in the limit test of OECD TG 203. The authors report for the dataset with the most precise values, the following distribution regarding species sensitivity: 51% algae most sensitive, 28.6% daphnia most sensitive, 18.6 % fish most sensitive; equal sensitivity in 1.2% of the cases; algae and daphnia equally sensitive 0.8%.
Hoekzema, C. C., A. J. Murk, B. J. v. d. Waart, J. C. M. v. d. Hoeven and D. F. d. Roode (2006). Alternative approaches can greatly reduce the number of fish used for acute toxicity testing. <i>Environmental Toxicology and Chemistry</i> 25(5): 1322-1325.	The authors applied the acute threshold (step-down) approach to data sets for 507 compounds, including agrochemicals, industrial chemicals, and pharmaceuticals from the internal NOTOX database. Theoretical applications of the acute threshold (step-down) approach gave similar results to those obtained with the standard fish median lethal concentration (LC50) test but required only 12% of the fish (3,195 instead of 27,324 fish

	used for all compounds in the database). In 188 (90%) of the 208 cases for which a complete data set was available, the EC50 for algae or daphnids was lower than the LC50 for fish. These results show that replacement of the standard fish LC50 test by the acute threshold (step-down) approach would greatly reduce the number of fish needed for acute ecotoxicity testing without any loss of reliability.
<i>Reduction / Refinement - general</i>	
Rufli, H. and T. A. Springer (2011). Can we reduce the number of fish in the OECD acute toxicity test? <i>Environmental Toxicology and Chemistry</i> 30(4): 1006-1011.	The authors evaluated acute fish toxicity data retrieved from an industrial database and the U.S. EPA Office of Pesticide Programs Ecotoxicity database and simulated various scenarios to improve the design of acute fish toxicity tests according to OECD TG 203. They concluded that the number of fish/concentration could be reduced to six without loss of statistical power. Moreover, they proposed to use the 96 h fish embryo acute toxicity test (now OECD TG 236) to define the concentrations for OECD TG 203. They further propose to use the 96 h fish embryo acute toxicity test in the step-down approach and confirm the concentration where no mortality occurs with a test using juvenile fish.
Rufli, H. (2012). Introduction of moribund category to OECD fish acute test and its effect on suffering and LC50 values. <i>Environmental Toxicology and Chemistry</i> 31(5): 1107-1112.	The term "moribund" is used to describe a state close to death. This state involves severe suffering and, in the interest of animal welfare, moribund animals should be removed from tests and humanely killed. The author uses a combination of severe sublethal symptoms (swimming behaviour, loss of equilibrium, ventilation, pigmentation) to define five moribund categories. Retrospective analysis of more than 400 acute fish toxicity studies revealed that a significant percentage (10-23%) of the fish in 49-79% of the studies would fulfil the moribund criteria, and in consequence, applying the criterion as endpoint would have reduced suffering of severity grade 3 (= severe distress) by up to 92 h. On the other side, the LC50 would be lower (by factor 2 in most cases) in 36-52% of the studies. Rufli recommends further research into the topic and clear definition of the moribund stage before introducing it into OECD TG 203.
Oris, J. T., S. E. Belanger and A. J. Bailer (2012). Baseline characteristics and statistical implications for the OECD TG 210 fish early-life stage chronic toxicity test. <i>Environmental Toxicology and Chemistry</i> 31(2): 370-376.	The authors systematically analysed the data and results of fish early-life stage tests (OECD TG 210) conducted in industry laboratories. The distribution of responses observed in control treatments was evaluated with the goal of understanding the implication of this variability on the sensitivity of OECD TG 210. The following recommendations on revised experimental design were given: maximise the number of replicate chambers per treatment concentration, increase the acceptable level of control hatching success and larval survival compared to current levels, using wet weight measurements rather than dry weight, and focusing test efforts on species that demonstrate less variability in outcome measures. From these analyses, the authors provide evidence of the impact of expected levels of variability on the sensitivity of traditional OECD TG 210 studies and the implications for defining a target for

	future animal alternative methods for chronic toxicity testing in fish.
<i>Cytotoxicity assays based on fish cell lines</i>	
Castaño, A., N. Bols, T. Braunbeck, P. Dierickx, M. Halder, B. Isomaa, K. Kawahara, L. Lee, C. Mothersill, P. Pärt, G. Repetto, J. Sintes, H. Rufli, R. Smith, C. Wood and H. Segner (2003). The use of fish cells in ecotoxicology. The report and recommendations of ECVAM Workshop 47. ATLA Alternatives to Laboratory Animals 31(3): 317-351.	The main objectives of this ECVAM workshop report were: - To critically review the potential use of fish cell to replace, reduce or refine existing regulatory tests which involve the use of fish for ecotoxicological purpose - To discuss the advantage, limitations and possible future applications of fish cells systems in hazard assessment, ecotoxicological research and testing and environmental surveillance and monitoring.
Bols, N. C., V. R. Dayeh, L. E. J. Lee and K. Schirmer (2005). Chapter 2 Use of fish cell lines in the toxicology and ecotoxicology of fish. Piscine cell lines in environmental toxicology. Biochemistry and Molecular Biology of Fishes. T. P. Mommsen and T. W. Moon, Elsevier. Volume 6: 43-84.	The authors review the use of fish cell lines in toxicology and ecotoxicology. Existing fish cell lines and their advantages, as well as comparison of fish cell lines, fish primary cultures and mammalian cell lines are covered. The authors report on general cellular responses (cytotoxicity, cell growth, genotoxicity and xenobiotic metabolism) studied on these models and highlight the toxicology (e.g. receptor-mediated toxicity, xenobiotic metabolism and cytotoxicity or mechanism of cell death) and ecotoxicology use of fish cell lines (e.g. ranking compounds for their potency, or assessing environmental samples, studying interactions between ecotoxicants and physical environment, developing and improving of biomarkers).
Ankley et al. (2010)	<i>See above – Mechanistic understanding of the endpoint</i>
Lakra, W. S., T. R. Swaminathan and K. P. Joy (2011). Development, characterization, conservation and storage of fish cell lines: a review. Fish Physiology and Biochemistry 37(1): 1-20.	This review gives an overview on teleost fish cells lines (their origins, development, characterization, conservation and storage) and provides a scientific update on new cell lines. It also discusses the importance of authentication, applications, cross-contamination and implications of overpassaged cell lines.
Lee, L. E. J., V. R. Dayeh, K. Schirmer and N. Bols (2009). Applications and potential uses of fish gill cell lines: examples with RTgill-W1. In Vitro Cellular and Developmental Biology - Animal 45(3-4): 127-134.	This work focuses on the importance of gills in fish biology and reviews the potential uses and applications of the RTgill-W1 fish cell lines. Their main advantages are: they can be grown in regular tissue culture surfaces or in transwell membranes in direct contact with water on their apical surfaces and environmental samples can be directly evaluated on these cells; they can withstand hypo- and hyper osmotic conditions; they have an optimal growth capacity at room temperature. All these features make them ideal sentinel models for <i>in vitro</i> aquatic toxicology and RTgill-W1 has been used to evaluated the toxicity of industrial effluent, polycyclic aromatic hydrocarbons, metals, or polybrominated diphenyl ethers (PBDEs); moreover, they are also used to study fish gill function and gills diseases.
Schirmer, K. (2006). Proposal to improve vertebrate cell cultures to establish them as substitutes for the regulatory testing of chemicals and effluents using fish. Toxicology 224(3): 163-183.	This review explains how <i>in vitro</i> assays based on vertebrate cell cultures could be improved so that a replacement of acute fish tests could more likely be achieved. The author reports that fish cell lines compare well with fish lethality in their relative sensitivity towards the toxicity of chemicals, as shown by strong correlation found for cell line <i>versus</i> fish acute toxicity test, excepted for organophosphate pesticides. This illustrates the first limit of these

	<p>models: a single culture cell will always have a limited number of target sites in comparison to whole organisms. The author also shows that fish cell lines are usually from one to several orders of magnitude less sensitive than whole organisms, which is the second limit of these models. According to the author, the diversity of target sites issue would have to be tackled by the establishment and/or characterization of cell culture models with functions typical of their origin. Concerning the sensitivity issue, various ways of improvement are highlighted, including the selection of appropriate cell lines, culture environment, endpoint measurements and means to account for the <i>in vitro</i> bioavailability of chemicals.</p>
<p>Fent, K. (2007). Permanent fish cell cultures as important tools in ecotoxicology. ALTEX: Alternativen zu Tierexperimenten 24 Spec No: 26-28.</p>	<p>This paper summarises some previous work in fish cell lines and highlights the importance of such <i>in vitro</i> systems for the acute toxicity assessment of a variety of environmental pollutants (organotins, substituted phenols, pharmaceuticals). It reports a significant correlation between cytotoxicity and physical-chemicals properties of the compounds on one hand, and between <i>in vitro</i> cytotoxicity data and fish acute toxicity (organotins, substituted phenols) and <i>Daphnia magna</i> (pharmaceuticals) on the other hand.</p>
<p>CEFIC LRI project - ECO 8 Development of a strategy to predict acute fish lethality using fish cell lines and fish embryos http://www.cefic-lri.org/projects/1224841821/21/ECO8-Development-of-a-strategy-to-predict-acute-fish-lethality-using-fish-cell-lines-and-fish-embryos/?cntnt01template=display_list_test</p>	<p>The goal of this research project was to investigate if <i>in vitro</i> approaches based on fish cell lines and/or fish embryos can be improved to be widely accepted as an alternative to the acute fish test. With regard to the cell lines, four issues were addressed in an attempt to overcome the seemingly lower absolute sensitivity. Firstly, several fish cell lines possessing origin-specific differentiated functions were employed. Secondly, the cell culture environment during toxicity testing was altered to increase the sensitivity of cells. Thirdly, a variety of endpoints were used in order to better reflect different modes of toxic action. Finally, the truly bio-available fraction of each chemical, rather than the nominally added concentration, was taken into account for concentration-response analysis. This latter issue will also be addressed for the same set of chemicals in the DarT.</p>
<p>Hestermann, E. V., J. J. Stegeman and M. E. Hahn (2000). Cell culture serum alters the uptake and relative potencies of halogenated aromatic hydrocarbons in PLHC-1 cells. Marine Environmental Research 50(1-5): 545-546.</p>	<p>This study focuses on the effect of serum in cell culture medium on the bioavailability of cytochrome P450 1A (CYP1A)-inducing compounds, in PLHC-1 cells (<i>Poeciliopsis lucida</i> hepatocellular carcinoma). The presence of 10% calf serum in the medium increased the EC50 for induction of ethoxyresorufin <i>O</i>-deethylase (EROD) activity by 2,3,7,8-tetrachlorodibenzo-<i>p</i>-dioxin (TCDD) 20-fold as compared to treatment in serum-free medium. Measurement of [3H]TCDD uptake and Ah receptor binding indicated that the apparent difference in potencies was a result of decreased bioavailability in the presence of serum, effectively reducing the concentration of TCDD within the cells. Induction of EROD and CYP1A protein in response to treatment with each of three coplanar polychlorinated biphenyls (PCB congeners 77, 126, and 169) was similarly affected by serum, although the magnitude varied among inducers and assays. Relative potencies (calculated as EC50TCDD /</p>

	<p>EC50PCB) for EROD induction by the three PCBs were significantly higher in the absence of serum. However, serum showed no significant effect on the relative potencies for CYP1A protein induction. These results demonstrate that measured inducing potencies, and relative potencies for EROD induction, by halogenated aromatic hydrocarbons are strongly dependent on the composition of culture medium, which can lead to artificial differences in comparisons among cell types.</p>
<p>Gülden, M. and H. Seibert (2005). Impact of bioavailability on the correlation between <i>in vitro</i> cytotoxic and <i>in vivo</i> acute fish toxic concentrations of chemicals. <i>Aquatic Toxicology</i> 72(4): 327-337.</p>	<p>The main objective of this study was to investigate whether a reduced availability of chemicals <i>in vitro</i> can account for the lower sensitivity of <i>in vitro</i> toxicity test systems. For this purpose, the bioavailable free fractions of the nominal cytotoxic concentrations (EC50) of chemicals determined with a cytotoxicity test system using the mammalian cell line Balb/c 3T3 and the corresponding free cytotoxic concentrations (ECu50) were calculated. The algorithm applied is based on a simple equilibrium distribution model for chemicals in cell cultures with serum-supplemented culture media, considering the distribution of chemicals between water, lipids and serum albumin.</p> <p>Organic chemicals covering a wide range of cytotoxic potency and lipophilicity were selected, for which fish acute toxicity data (LC50-values) from at least one of the three fish species, medaka, rainbow trout and fathead minnow, were available. The availability of several chemicals was shown to be extensively reduced either by partitioning into lipids or by serum albumin binding or due to both mechanisms, and reduction of bioavailability became more important with increasing cytotoxic potency. The sensitivity of the Balb/c 3T3 cytotoxicity assay and the correspondence between <i>in vivo</i> and <i>in vitro</i> toxic potencies were increased when the free cytotoxic concentrations instead of the nominal cytotoxic concentrations were used. The few remaining differences between cytotoxic and acute toxic concentrations can be explained by a more specific mechanism of acute toxic action than basal cytotoxicity. It is concluded that the frequently observed low sensitivity of <i>in vitro</i> cytotoxicity test systems, compared to fish acute toxicity assays, at least in part, can be explained by differences in the availability of chemicals <i>in vitro</i> and <i>in vivo</i>. Moreover, neglecting these differences systematically causes a bias of the correlation between <i>in vivo</i> and <i>in vitro</i> toxic potencies of chemicals, whereas taking them into account increases the predictivity of the <i>in vitro</i> assays.</p>
<p>Tanneberger, K., A. Rico-Rico, N. I. Kramer, F. J. M. Busser, J. L. M. Hermens and K. Schirmer (2010). Effects of Solvents and Dosing Procedure on Chemical Toxicity in Cell-Based <i>in Vitro</i> Assays. <i>Environmental Science & Technology</i> 44(12): 4775-4781.</p>	<p>The aims of this study were to quantify the exposure and associated toxicity of chemicals with different physicochemical properties toward a fish gill cell line when different solvents and procedural steps are used to introduce test chemicals to cells. Three chemicals with a range of hydrophobicity and volatility were selected and delivered in three different solvents using two common dosing procedures. Toxicity tests were coupled with chemical analysis to</p>

	<p>quantify the chemical concentrations within culture wells. The results show that the impact of solvents and dosing procedure was greatest for the most volatile and hydrophobic test chemical, and that certain combinations of the test chemical, solvent, and procedural steps can lead to inhomogeneous distribution of the test chemical and thus differing degrees of bioavailability, resulting in quantitative differences in apparent toxicity.</p>
<p>Tanneberger, K., M. Knöbel, F. J. M. Busser, T. L. Sinnige, J. L. M. Hermens and K. Schirmer (2013). Predicting fish acute toxicity using a fish gill cell line-based toxicity assay. <i>Environmental Science & Technology</i> 47: 1110-1119.</p>	<p>This article summarises the outcome of the CELLsens project. The authors developed a fish gill cell line-based (RTgill-W1) assay, using several measures to improve sensitivity. The optimized assay was applied to determine the toxicity of 35 organic chemicals, having a wide range of toxicity to fish, mode of action and physicochemical properties. The authors found a very good agreement between <i>in vivo</i> and <i>in vitro</i> effective concentrations. For up to 73% of the tested compounds, the difference between the two approaches was less than 5-fold, covering baseline toxicants but as well compounds with presumed specific modes of action, including reactivity, inhibition of acetylcholine esterase or uncoupling of oxidative phosphorylation. Accounting for measured chemical concentrations eliminated two outliers (one hydrophobic and one volatile). Few outliers remained. The most striking were allyl alcohol (2700-fold), which likely needs to be metabolically activated; permethrin (190-fold), and lindane (63-fold), compounds acting, respectively, on sodium and chloride channels in the brain of fish. The authors also discuss further developments of this assay and suggest its use beyond predicting acute toxicity to fish, for example, as part of adverse outcome pathways to replace, reduce, or refine chronic fish tests.</p>
<p>Schirmer, K., A. G. J. Chan, B. M. Greenberg, D. G. Dixon and N. C. Bols (1997). Methodology for demonstrating and measuring the photocytotoxicity of fluoranthene to fish cells in culture. <i>Toxicology in Vitro</i> 11(1–2): 107-119.</p>	<p>The authors have developed a methodology for quantifying the photocytotoxicity of fluoranthene to a gill cell line from rainbow trout. Solubilisation in a modified culture medium was achieved with and without foetal bovine serum (FBS) and with and without dimethyl sulfoxide (DMSO). FBS caused most of the fluoranthene to remain in solution and blocked photocytotoxicity if present during UV irradiation. DMSO had little effect on fluoranthene distribution in cell cultures but caused cells to be slightly more sensitive to the phototoxicity of fluoranthene, which indicate that both the presence of FBS and the type of solvent can influence the outcome of the assay. The indicator dyes alamar blue and 5-carboxyfluorescein diacetate acetoxymethyl ester were used to quantify cytotoxicity in two different ways: singly in two separate assays, and mixed together. With UV irradiation for 2 h, both dyes indicated increasing loss of viability with increasing doses of fluoranthene. EC50 values ranged from 18 to 44 ng/ml (89–217 nM), with the alamar blue assay being slightly more sensitive.</p>
<p>Ackermann, G. E. and K. Fent (1998). The adaptation of the permanent fish cell lines PLHC-1 and RTG-2 to FCS-free media</p>	<p>The omission of foetal calf serum (FCS) from culture medium can influence the response of cells in assays measuring cytotoxicity, cytochrome P450 induction or estrogenic activity of</p>

<p>results in similar growth rates compared to FCS-containing conditions. <i>Marine Environmental Research</i> 46(1–5): 363-367.</p>	<p>chemicals. The purpose of this study was to provide a selection of different FCS-free cell culture media in which the two established fish cell lines PLHC-1 (a hepatoma cell line derived from <i>Poeciliopsis lucida</i>) and RTG-2 (a gonadal cell line derived from <i>Oncorhynchus mykiss</i>) grow as well as in FCS-containing medium. The results shows that PLHC-1 and RTG-2, adapted to FCS-free conditions, show similar growth rates as when thriving in FCS-containing medium. Cultivation under FCS-free conditions may represent an improvement in the widely used cytotoxicity tests or cytochrome P450-activity assays and is a prerequisite of <i>in vitro</i> systems for detecting estrogenic compounds in fish.</p>
<p>Schirmer, K., K. Tanneberger, N. I. Kramer, D. Völker, S. Scholz, C. Hafner, L. E. J. Lee, N. C. Bols and J. L. M. Hermens (2008). Developing a list of reference chemicals for testing alternatives to whole fish toxicity tests. <i>Aquatic Toxicology</i> 90(2): 128-137.</p>	<p>This paper details the derivation of a list of 60 reference chemicals for the development of alternatives to animal testing in ecotoxicology with a particular focus on fish, for the CELLSens project. This project aims to systematically gather mechanistic information on the performance of alternative testing methods. Specifically, the ability of fish cell lines and zebrafish embryos to detect specific modes of action, and the role of dosing and exposure schemes in the expression of toxicity of chemicals with a wide range of physico-chemical properties will be investigated. These investigations called for a carefully selected list of reference chemicals.</p> <p>The chemicals were selected as a prerequisite to gather mechanistic information on the performance of alternative testing systems (vertebrate cell lines and fish embryos), in comparison to the fish acute lethality test. The U.S. EPA fathead minnow database was consulted as reference for whole organism responses. This database was compared to the Halle Registry of Cytotoxicity and a collation of data by the German EPA (UBA) on acute toxicity data derived from zebrafish embryos. Chemicals that were present in the fathead minnow database and in at least one of the other two databases were subject to selection. Criteria included the coverage of a wide range of toxicity and physico-chemical parameters as well as the determination of outliers of the <i>in vivo/in vitro</i> correlations.</p> <p>According to the authors, this reference list could also be of benefit to search for alternatives in ecotoxicology in general, as a reference set of chemicals whose common use could help accelerate the development of non-animal alternatives in toxicology and ecotoxicology. One example would be the use of this list to validate structure–activity prediction models, which in turn would benefit from a continuous extension of this list with regard to physico-chemical and toxicological data.</p>
<p>Knauer, K., C. Lampert and J. Gonzalez-Valero (2007). Comparison of <i>in vitro</i> and <i>in vivo</i> acute fish toxicity in relation to toxicant mode of action. <i>Chemosphere</i> 68(8): 1435-1441.</p>	<p>The authors investigated the acute toxicity of 18 plant protection products to the fish hepatoma cell line PLHC-1 and to juvenile rainbow trout, the main objective being to explore whether hepatoma cells could be used to predict acute toxicity in fish taking into account the mode of toxic action and compound properties. Acute fish toxicity was determined using the</p>

	<p>OECD guideline test 203 and compared to predicted baseline LC50 of acute fish toxicity calculated with a quantitative structure–activity relationship (QSAR) derived for guppy fish. Cytotoxicity was determined through the inhibition of neutral red uptake (NR50) into lysosomes and compared to predicted baseline cytotoxicity derived for goldfish GFS cells. In general, NR50 values were higher by a factor ranging from 3 to 3000 than the corresponding acute LC50. A weak correlation between NR50 and LC50 values was found (log/log: $r^2 = 0.62$). The lipophilicity (log Kow) was not a good predictor for cytotoxicity ($r^2 = 0.43$) and lethality ($r^2 = 0.57$) of these pesticides. The neutral red assay is detecting general baseline toxicity only, and may be of a limited use in predicting acute fish toxicity.</p> <p>Comparing LC50 data to QSAR results, the compounds can be classified to act as narcotics or reactive compounds with a specific mode of toxic action in fish. According to the authors, a set of <i>in vitro</i> tests investigating effects on different endpoints in various cell culture models might be a suitable replacement for acute <i>in vivo</i> toxicity tests with fish, the challenge for selecting functional endpoints being to identify those that do have predictive value for the expression of toxicity.</p>
<p>Tollefsen, K. E., C. Blikstad, S. Eikvar, E. Farnen Finne and I. Katharina Gregersen (2008). Cytotoxicity of alkylphenols and alkylated non-phenolics in a primary culture of rainbow trout (<i>Onchorhynchus mykiss</i>) hepatocytes. <i>Ecotoxicology and Environmental Safety</i> 69(1): 64-73.</p>	<p>The aim of this study was to assess the cytotoxicity of a range of alkylphenols and alkylated non-phenolics in a primary culture of rainbow trout (<i>Onchorhynchus mykiss</i>) hepatocytes to construct a structure–toxicity relationship for this group of ubiquitous aquatic pollutants. The metabolic inhibition and loss of membrane integrity were used as cytotoxic endpoints through use of the cellular markers alamar blue and 5 carboxyfluorescein diacetate acetoxyethyl ester, respectively. The results show that cytotoxicity increased with the hydrophobicity of the alkylphenols for compounds with log Kow < 4.9. Normal chained alkylphenols, branched alkylphenols and multi-substituted alkylphenols with log Kow ≥ 4.9 deviated clearly from this relationship. The alkylphenols displayed greater cytotoxicity than alkylated non-phenolics and it is proposed that most alkylated non-phenolic caused non-polar narcosis (baseline toxicity) whereas the alkylphenols caused polar narcosis. Observations that metabolic inhibition occurred at lower concentrations than loss of membrane integrity for most chemicals indicated that interference with cellular metabolic functions was the main cause of cytotoxicity. Metabolic inhibition corresponded better than loss of membrane integrity to reported acute toxicity to fish, although the <i>in vivo</i> acute toxicity of hydrophobic compounds (log Kow > 2-3) was clearly underestimated by both endpoints.</p>
<p>Kramer, N. I., J. L. M. Hermens and K. Schirmer (2009). The influence of modes of action and physicochemical properties of chemicals on the correlation between <i>in vitro</i> and acute fish toxicity data. <i>Toxicology in Vitro</i> 23(7): 1372-1379.</p>	<p>The objective of this study was to determine what factors influence the correlation between <i>in vitro</i> and fish toxicity data. Basal cytotoxicity (IC50) from mammalian cells and acute toxicity data from fathead minnow (LC50) of 82 industrial organic chemicals were obtained from the Halle Registry of Cytotoxicity and the US EPA Fathead Minnow Database and</p>

	<p>compared. A good correlation between IC50 with LC50 data was found ($r = 0.84$). Yet, IC50 data were less sensitive than LC50 data by an order of magnitude.</p> <p>Using multiple regression analysis, the octanol–water partition coefficient ($\log K_{ow}$) and the Henry’s Law Constant (H) were found to significantly explain the low absolute sensitivity. The mode of action (MOA) of the chemical was found to significantly explain the general variation in the $\log IC_{50}/\log LC_{50}$ regression line. These results support the notion that (a) the bioavailability of hydrophobic (high $\log K_{ow}$) and volatile (high H) chemicals is significantly lower in <i>in vitro</i> assays than in the fish bioassay and (b) multiple cell types and endpoints should be included to mimic the modes of action possible in the whole organism. Moreover, the free concentration is a better measure of exposure in <i>in vitro</i> assays and may be modelled using a compound’s $\log K_{ow}$ and H.</p>
<p>Segner, H., A. Behrens, E. M. Joyce, K. Schirmer and N. C. Bols (2000). Transient induction of 7-ethoxyresorufin-O-deethylase (EROD) activity by medium change in the rainbow trout liver cell line, RTL-W1. <i>Marine Environmental Research</i> 50(1–5): 489-493.</p>	<p>Cytochrome P4501A (CYP1A) metabolises a wide array of lipophilic xenobiotics. In fish liver, CYP1A is constitutively expressed at low levels, and xenobiotics can strongly induce CYP1A expression via a receptor-mediated pathway. The authors show that in the rainbow trout liver cell line, RTL-W1, CYP1A-catalysed 7-ethoxyresorufin-O-deethylase (EROD) activity can be induced by a change of the culture medium, in the absence of xenobiotics. The increase in cellular EROD levels is of transient nature, and the results obtained indicate that photooxidised tryptophan is the agent causing the increase of EROD activity after medium change.</p>
<p>Fent, K. (2001). Fish cell lines as versatile tools in ecotoxicology: assessment of cytotoxicity, cytochrome P4501A induction potential and estrogenic activity of chemicals and environmental samples. <i>Toxicology in Vitro</i> 15(4–5): 477-488.</p>	<p>This paper illustrates the versatility and high potential of fish cell lines in ecotoxicology. It presents fish cell lines and their use in cytotoxicity testing, and highlights the importance of aryl hydrocarbons receptor (AHR) and cytochrome P450 induction and dioxin-like activity in fish cell lines and in the ecotoxicology of fish. It presents the TEQ (toxic equivalency) concept used to estimate the toxic potential of environmental mixtures of halogenated aromatic hydrocarbons. It also reports the use of fish-related <i>in vitro</i> system for the detection of hormonally active compounds such as xeno-estrogens, and the development of a RTG-2 cell-based reporter gene system, based on the transfection of a reporter gene plasmid consisting of an estrogen responsive element (ERE) fused to a firefly luciferase gene, which induces luciferase expression after binding of an estrogenic agonist to the estrogen receptor (ER) and transcriptional activation. The use of fish cell lines as bioassays for evaluation of environmental samples contaminated with CYP1A inducing compounds or estrogenic compounds is discussed.</p>
<p>Hestermann, E. V., J. J. Stegeman and M. E. Hahn (2002). Serum withdrawal leads to reduced aryl hydrocarbon receptor expression and loss of cytochrome P4501A inducibility in PLHC-1 cells.</p>	<p>The authors tested the ability of changes in serum used in cell culture medium to alter expression of the AHR and induction of cytochrome P4501A (CYP1A) in PLHC-1 cells. Culture of early-passage cells in serum-free medium for 2 days led to a loss of CYP1A</p>

<p>Biochemical Pharmacology 63(8): 1405-1414.</p>	<p>inducibility by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), whereas culture in 10% delipidated calf serum increase the TCDD-induced levels of both CYP1A protein and enzymatic activity relative to levels in cells cultured in 10% complete calf serum. The kinetics of induction were unaffected.</p> <p>In cells cultured in serum-free medium for 1 and 2 days there was a progressive loss of CYP1A inducibility and a time-dependent decline in AHR protein. The measured reduction in AHR could be shown to predict the loss of CYP1A induction. Expression of AHR protein was unaffected by culture in 10% delipidated serum. The effects of serum-free medium and delipidated were unaffected by serum withdrawal.</p> <p>Comparison of early- and late-passage cells revealed a twofold greater rate of proliferation in the latter, suggesting that a growth advantage is coincident with loss of the serum-dependency of AHR expression. These results provide a quantitative link between changes in receptor expression and a downstream response, establishing a foundation for future studies of receptors expression and sensitivity to toxic responses <i>in vitro</i> and <i>in vivo</i>.</p>
<p>Bopp, S. K., N. Bols and K. Schirmer (2006). Development of a solvent-free, solid-phase in vitro bioassay using vertebrate cells. Environmental Toxicology and Chemistry 25(5): 1390-1398.</p>	<p>One problem common to miniaturized systems is the loss of test chemicals because of sorption. In the present study, the authors used the sorption phenomenon in a positive way: as it was found that contaminants sorbed to the growth surface in wells of tissue-culture plates are available to vertebrate cells growing in direct contact with the contaminant-coated surface, they used Biosilon, a bead cell-culture carrier made of polystyrene, and contaminated with polycyclic aromatic hydrocarbons (PAHs), to support cell adherence. This allowed the detection of reproducible dose–response curves of an increase in cytochrome CYP1A enzyme activity by sorbed PAHs in the rainbow trout (<i>Oncorhynchus mykiss</i>) liver cell line, RTL-W1. The resulting bead assay provides a miniaturized, solvent-free exposure system. Potential future applications include the coupling to environmental sampling, in which the bead material is used as solid receiving phase before serving as a surface for vertebrate cells to attach and respond.</p>
<p>Kramer, N. I., F. J. M. Busser, T. T. Oosterwijk, K. Schirmer, B. Escher and J. L. M. Hermens (2010). Development of a partition-controlled dosing system for cell assays. Chemical Research in Toxicology 23: 1806-1814.</p>	<p>Hydrophobic and volatile chemicals have proven to be difficult to dose in cell assays, and the free concentration of these chemicals in culture medium may diminish over time due to metabolism, evaporation, and nonspecific binding to well plate surfaces and serum constituents. The aim of this study was therefore to develop a partition controlled dosing system to maintain constant concentrations of benzo(a)pyrene, 1,2-dichlorobenzene, and 1,2,4-trichlorobenzene in an ethoxyresorufin-<i>O</i>-deethylase (EROD) assay and a cytotoxicity assay with the rainbow trout (<i>Oncorhynchus mykiss</i>) cell lines RTL-W1 and RTgill-W1. Polydimethylsiloxane (PDMS) sheets were loaded with test chemicals in a spiked methanol/water solution and placed in the wells, filled with culture medium, of a 24-well</p>

	<p>culture plate. Cells were grown on inserts and were subsequently added to the wells with the PDMS sheets. The system reached equilibrium within 24 h. The reservoir of test chemical in PDMS allow to compensate for the loss of >95% of the test chemical from the culture medium. The PDMS sheets maintained medium concentrations constant for >72 h. Nominal median effect concentrations (EC50) were 1.3-7.0 times lower in the partition-controlled dosing systems than in conventional assays spiked using dimethyl sulfoxide (DMSO) as a carrier solvent, thus indicating that the apparent sensitivity of the bioassay increased when controlled and constant exposure conditions could be assured. The EC50 values of the test chemicals based on free concentrations were estimated in the partition-controlled dosing systems using measured PDMS-bare culture medium partition coefficients. Results indicated that 61, 70, and 99.8% of 1,2-diclorobenzene, 1,2,4-trichlorobenzene, and benzo(a)pyrene were bound to serum constituents in the culture medium.</p>
<p><i>Adverse Outcome Pathways</i></p> <p>Volz, D. C., S. Belanger, M. Embry, S. Padilla, H. Sanderson, K. Schirmer, S. Scholz and D. Villeneuve (2011). Adverse Outcome Pathways during Early Fish Development: A Conceptual Framework for Identification of Chemical Screening and Prioritization Strategies. <i>Toxicological Sciences</i> 123(2): 349-358.</p>	<p>The fish early life-stage (FELS) test guideline (OECD TG 210) is the most frequently used bioassay for predicting chronic fish toxicity and supporting aquatic ecological risk assessments. Although valuable for predicting fish full life-cycle toxicity, this test is labor and resource intensive and, due to an emphasis on apical endpoints, provide little to no information about chemical mode of action. Therefore, the development and implementation of alternative testing strategies for screening and prioritizing chemicals could reduce the cost and number of animals required for estimating FELS toxicity and, at the same time, provides insights into mechanisms of toxicity. In this context, the authors proposed 3 FELS-specific adverse outcome pathways (AOPs) as conceptual frameworks for identifying useful chemical screening and prioritization strategies, using three reference chemicals: a cardiotoxic aryl hydrocarbon receptor agonist (2,3,7,8-tetrachlorodibenzo-p-dioxin), a neurotoxic acetylcholinesterase inhibitor (chlorpyrifos), and a narcotic surfactant (linear alkylbenzene sulfonate). Using qualitative descriptions for each chemical during early fish development, they developed generalized AOPs and, based on these examples, proposed a three-tiered testing strategy for screening and prioritizing chemicals for FELS testing. According to the authors, this tiered testing strategy linked with biologically based concentration-response models, could help reduce the reliance on long-term and costly FELS tests required for assessing the hazard of thousands of chemicals currently in commerce.</p>
<p>Villeneuve, D., Volz, D.C., Embry, M.R., Ankley, G., Belanger, S.E., Léonard, M., Schirmer, K., Tanguay, R., Truong, L, and Wehmas, L. (in preparation) Alternatives to the Fish Early Life-Stage Test: A Strategy for Discovering and Annotating Adverse</p>	<p>This is the report of an ILSI HESI expert workshop organised in 2012. Participants discussed the development of efficient and cost-effective alternatives to the fish early life-stage (FELS) test (OECD TG 210) used to estimate chronic fish toxicity in support of ecological risk assessments and global chemical management programs. Identification and description of</p>

<p>Outcome Pathways for Early Fish Development.</p>	<p>potential adverse outcome pathways (AOPs) relevant to FELS toxicity would support the development of such alternatives. An overall strategy for discovery and annotation of FELS AOPs was outlined and key events represented by major developmental landmarks were organised into a preliminary conceptual model of fish development. Using swim bladder inflation as an example, a weight-of-evidence-based approach was used to support linkage of key molecular initiating events to adverse phenotypic outcomes and reduced young-of-year survival. Based on an iterative approach, the feasibility of using key events as the foundation for expanding a network of plausible linkages and AOP knowledge was explored, important knowledge gaps identified, and key research objectives were defined. The example and strategy described are intended to guide collective efforts to define FELS-related AOPs and develop resource-efficient predictive assays that address the toxicological domain of the OECD TG 210 test.</p>
<p>CEFIC LRI funded project (LRI-ECO20-UA) <i>Development of an alternative testing strategy for the fish early life-stage test for predicting chronic toxicity</i> http://www.cefic-lri.org/projects/35/21/LRI-ECO20-UA-Development-of-an-alternative-testing-strategy-for-the-fish-early-life-stage-test-for-predicting-chronic-toxicity/?cntnt01template=display_list_test</p>	<p>The main objective of this project is to use a mechanistic framework to develop and propose a high-throughput tiered-testing strategy for screening and prioritizing chemicals for FELS testing (OECD 210). The main objective can be divided into four parts:</p> <ol style="list-style-type: none"> 1) Establish a database of toxicologically relevant FELS-specific AOPs, identify molecular initiating events and subsequent intermediate responses resulting into the apical outcome of interest 2) Propose Tier 1 in vitro screening toolbox to test for AOP-specific events and responses predictive for FELS chronic toxicity. 3) Propose Tier 2 whole-organism ZFET assays to test for AOP-specific events and responses predictive for FELS chronic toxicity and assess the potential of a ZFET molecular screening tool to predict cellular, organ and/or organism responses giving rise to FELS chronic toxicity. 4) Offering a proposal for implementation of a tiered-testing strategy in EU regulation. Assessment of usefulness and applicability of tiered testing strategy for global scientific and regulatory community
<p>OECD (2011). Report of the workshop on using mechanistic information in forming chemical categories. Series on Testing and Assessment n° 138. ENV/JM/MONO(2011)8; OECD, Paris, France.</p>	<p>This document is a report of the Workshop on Using Mechanistic Information in Forming Chemical Categories, held on 8-10 December 2010 in Crystal City VA, USA. One of the aims of the OECD QSAR project is to build tools which allow the user to fill data gaps by using existing data for similar chemicals. Using mechanistic characteristics to group similar chemicals has been shown to be very successful with the OECD QSAR Toolbox for some endpoints such as skin sensitisation or acute aquatic toxicity. A key aspect in forming toxicologically meaningful categories is identifying mechanistic characteristics (i.e. key events and processes) which relate to the risk assessment endpoint in question and can be</p>

	<p>measured or predicted.</p> <p>The purpose of this workshop was to acquire scientific input which will guide further development and use of the concept of the adverse outcome pathway (AOP). An AOP delineates the documented, plausible, and testable processes by which a chemical induces molecular perturbations and the associated biological responses which describe how the molecular perturbations cause effects at the subcellular, cellular, tissue, organ, whole animal and (when required) population levels of observation.</p> <p>The aims of the workshop were to:</p> <ul style="list-style-type: none"> - Review the extent of the knowledge on mechanism or mode of action in the context of key events or processes that lead to specific adverse outcomes that are used in risk assessment; - Propose how scientific information on mechanism or mode of action can be organised as key events and processes within adverse outcome pathways to aid the formation of categories; - Examine a series of case studies using adverse outcome pathways; - Gather input on work flow(s) for using adverse outcome pathways to form chemical categories, <p>and to</p> <ul style="list-style-type: none"> - Gather input on the role of (Q)SAR methods in forming categories based on key events in an adverse outcome pathway. <p>The example of AOP presented in this document are the following:</p> <ul style="list-style-type: none"> - Weak acid respiratory uncoupling and acute aquatic toxicity - ER-mediated reproductive impairment and reproductive toxicity - Voltage gated sodium channels and neurotoxicity - Haemolytic anaemia induced by anilines and nephrotoxicity induced by 4-aminophenols and repeated dose toxicity <p>Using those AOP proposals, the authors also develop chemical categories corresponding to these modes of action.</p>
<p>Yozzo, K. L., S. P. McGee and D. C. Volz (2013). Adverse outcome pathways during zebrafish embryogenesis: A case study with paraoxon. <i>Aquatic Toxicology</i> 126(0): 346-354.</p>	<p>The objective of this study was to develop an adverse outcome pathway (AOP) that provided quantitative linkages across levels of biological organization during zebrafish embryogenesis, using paraoxon as a reference acetylcholinesterase (AChE) inhibitor.</p> <p>Within normal zebrafish embryos, the authors first demonstrated that ache transcripts and AChE activity increased in a stage-dependent manner following segmentation. They then showed that static exposure of embryos to paraoxon (31.2–500 nM) from 5 to 96 hpf resulted in significant stage- and concentration-dependent AChE inhibition, albeit these effects were fully reversible within 48 h following transfer to clean water. However, even in the presence</p>

	<p>of significant AChE inhibition, exposure to non-teratogenic paraoxon concentrations (≤ 250 nM) did not adversely impact secondary motoneuron development at 96 hpf. They also investigated the potential effects of paraoxon exposure on spontaneous tail contractions at 26 hpf – an early locomotor behaviour that results from innervation of primary motoneuron axons to target axial muscles, and found that the frequency of spontaneous tail contractions at 26 hpf – a developmental stage with minimal AChE expression and activity – was significantly higher following exposure to paraoxon concentrations as low as 31.2 nM. Those results suggest that (1) normal AChE activity is not required for secondary motoneuron development and (2) spontaneous tail contractions at 26 hpf are sensitive to paraoxon exposure, an effect that may be independent of AChE inhibition. Using a well-studied reference chemical, this study highlights the potential challenges in developing quantitative AOPs to support chemical screening and prioritization strategies.</p>
<p><i>Use of fish embryos for chronic toxicity testing</i></p> <p>Weil, M., S. Scholz, M. Zimmer, F. Sacher and K. Duis (2009). Gene expression analysis in zebrafish embryos: a potential approach to predict effect concentrations in the fish early life stage test. <i>Environmental Toxicology and Chemistry</i> 28(9): 1970-1978.</p>	<p>Based on the hypothesis that analysis of gene expression could be used to predict chronic fish toxicity, the authors used a 48 h zebrafish (<i>Danio rerio</i>) embryo test and expanded it to a gene expression <i>D. rerio</i> embryo test (Gene-DarT). The effects of 14 substances on lethal and sublethal endpoints of the DarT and on expression of potential marker genes were investigated: the aryl hydrocarbon receptor 2, cytochrome P450 1A (cyp1a), heat shock protein 70, fizzy-related protein 1, the transcription factors v-maf musculoaponeurotic fibrosarcoma oncogene family protein g (avian) 1 and NF-E2-p45-related factor, and heme oxygenase 1 (hmox1). After exposure of zebrafish embryos for 48 h, differential gene expression was evaluated using reverse transcriptase–polymerase chain reaction, gel electrophoresis, and densitometric analysis of the gels. All tested compounds significantly affected the expression of at least one potential marker gene, with cyp1a and hmox1 being most sensitive. Lowest observed-effect concentrations (LOECs) for gene expression were below concentrations resulting in 10% lethal effects in the DarT. For 10 (3,4- and 3,5-dichloroaniline, 1,4-dichlorobenzene, 2,4-dinitrophenol, atrazine, parathion-ethyl, chlorotoluron, genistein, 4-nitroquinoline-1-oxide, and cadmium) out of the 14 tested substances, LOEC values derived with the Gene-DarT differ by a factor of less than 10 from LOEC values of fish early life stage tests with zebrafish. For pentachloroaniline and pentachlorobenzene, the Gene-DarT showed a 23- and 153-fold higher sensitivity, respectively, while for lindane, it showed a 13-fold lower sensitivity. For ivermectin, the Gene-DarT was by a factor of more than 1,000 less sensitive than the acute fish test. The results of the present study indicate that gene expression analysis in zebrafish embryos could principally be used to predict effect concentrations in the fish early life stage test.</p>

<p>Species extrapolation</p> <p>Ahlers, J., C. Riedhammer, M. Vogliano, R.-U. Ebert, R. Kühne and G. Schüürmann (2006). Acute to chronic ratios in aquatic toxicity - Variation across trophic levels and relationship with chemical structure. <i>Environmental Toxicology and Chemistry</i> 25(11): 2937-2945.</p>	<p>The authors determined acute to chronic ratios (ACRs) for fish, daphnids, and algae, for new and existing chemicals. They used only test results derived in compliance with the European Union Technical Guidance Document (TGD) and validated by authorities. The results show that the median ACRs of 10.5 (fish), 7.0 (daphnids), and 5.4 (algae) are well below the ACR safety factor of 100 as implied by the TGD; however individual ACRs vary considerably and go up to 4,400, which suggest that a safety factor of 100 is not protective for all chemicals and trophic levels. They did neither find correlation between ACR and baseline toxicity nor between ACR across trophic levels. Narcosis is associated with a preference for a low ACR; nevertheless, low ACRs are frequently obtained for nonnarcotics. Analysis of chemical structures led to the derivation of structural alerts to identify compounds with a significantly increased potential for a high ACR, which may prove to be useful in setting test priorities. Thus, according to the authors, life-cycle tests are the only way to conservatively predict long-term toxicity.</p>
<p>Weyers, A., B. Sokull-Klüttgen, J. Baraibar-Fentanes and G. Vollmer (2000). Acute toxicity data: A comprehensive comparison of results of fish, daphnia, and algae tests with new substances notified in the European Union. <i>Environmental Toxicology and Chemistry</i> 19(7): 1931-1933.</p>	<p>Acute aquatic toxicity data (fish, <i>Daphnia</i>, and algae) from the New Chemicals Database of the European Chemicals Bureau were compared according to their sensitivity. The algal growth inhibition test was the most sensitive (43.5%). For 18.5% cases, species were equally sensitive. In 37.9% cases, either fish or <i>Daphnia</i> tests were more sensitive. Correlation between fish and <i>Daphnia</i> toxicity values was better ($r^2=0.597$) than the correlation with algae, which might reflect similarities in physiology between the two animals not shared by algae. Toxicity values among the three organisms correlated more strongly than any with log <i>Pow</i>.</p>
<p>Hutchinson et al. (2003)</p>	<p>See above REDUCTION - Step-down Approach</p>
<p>Jeram et al. (2005)</p>	<p>See above REDUCTION - Step-down Approach</p>
<p>Hoekzema et al. (2006)</p>	<p>See above REDUCTION - Step-down Approach</p>
<p>Tebby, C., E. Mombelli, P. Pandard and A. R. R. Péry (2011). Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish. <i>Science of the Total Environment</i> 409(18): 3334-3343.</p>	<p>In this paper, the authors analysed acute toxicity data gathered by the Japanese Ministry of Environment for three species belonging to three different trophic levels (i.e., <i>Pseudokirchneriella subcapitata</i> 72 h EC50, <i>Daphnia magna</i> 48h EC50 and <i>Oryzias latipes</i> 96 h LC50), and investigate the relationships between the chemical structure and both the toxicity of the chemicals and the cross-species differences in sensitivity. In predicting toxicity towards each species, simple linear regression on logP performed better than PLS regression of toxicity on a very large set of molecular descriptors; moreover the predictions based on the interspecies correlations performed better than the QSAR predictions, which, in terms of cross-species comparisons, encourage the use of test strategies focussing on</p>

	reducing the number of tests on fish.
Netzeva et al. (2007)	<i>See above - QSAR</i>
Zvinavashe, E., T. Du, T. Griff, H. H. J. v. d. Berg, A. E. M. F. Soffers, J. Vervoort, A. J. Murk and I. M. C. M. Rietjens (2009). Quantitative structure-activity relationship modeling of the toxicity of organothiophosphate pesticides to <i>Daphnia magna</i> and <i>Cyprinus carpio</i> . <i>Chemosphere</i> 75(11): 1531-1538.	<p>The objective of this study was to develop QSAR models to describe the acute toxicity of organothiophosphate pesticides to aquatic organisms. Literature data sets for acute toxicity data of organothiophosphates to fish and one data set from experiments with 15 organothiophosphates on <i>Daphnia magna</i> performed in the present study were used to establish QSARs.</p> <p>Additionally, the authors investigated if toxicity data for the invertebrate <i>D. magna</i> could be used to build a QSAR model to predict toxicity to fish. Suitable QSAR models ($0.80 < r^2 < 0.82$) were derived to predict acute toxicity of organothiophosphates to fish (<i>Cyprinus carpio</i>) and the invertebrate (<i>D. magna</i>). The results shows that toxicity data for <i>D. magna</i> correlated well ($r^2 = 0.94$) with toxicity data for <i>C. carpio</i>, which implies that by performing toxicity tests with <i>D. magna</i>, one can use this interspecies QSAR model to predict the acute toxicity of organothiophosphates to fish. Finally, from the 100196 European Inventory of Existing Commercial Chemical Substances (EINECS), 83 compounds were identified that fit the selection criteria for the QSAR models. By applying the QSAR models to these chemicals, the authors obtained an indication of their toxicity without the need for additional experimental testing.</p>
Kar, S. and K. Roy (2010). First report on interspecies quantitative correlation of ecotoxicity of pharmaceuticals. <i>Chemosphere</i> 81(6): 738-747.	<p>Pharmaceuticals are designed to have a specific mode of action and many of them are persistent in the body, which makes them of concern for potential effects on aquatic flora and fauna. As there is a general scarcity of publicly available ecotoxicological data concerning pharmaceuticals, interspecies toxicity correlations could provide a tool for estimating contaminant sensitivity with known levels of uncertainty for a diversity of wildlife species. In this context, the authors have developed interspecies toxicity correlation between <i>Daphnia magna</i> (zooplankton) and fish, assessing the ecotoxicological hazard potential of diverse 77 pharmaceuticals. The developed models are validated and a consensus models are presented to predict toxicity of the individual compounds for any one species when the data for the other species are available ($r^2=0.75$ and $r^2=0.724$ for the prediction of daphnia and fish toxicity respectively). Informative illustrations of the contributing structural fragments which are responsible for the greater toxicity of the diverse pharmaceuticals are also identified by the developed models, and the developed models are used to predict fish toxicities of 59 pharmaceuticals (from <i>Daphnia</i> toxicity data) and <i>Daphnia</i> toxicities of 30 pharmaceuticals (from fish toxicity data).</p>
Zhang, X. J., H. W. Qin, L. M. Su, W. C. Qin, M. Y. Zou, L. X. Sheng, Y. H. Zhao and M. H. Abraham (2010). Interspecies	This article describes the link between the interspecies relationship and the differences of bio-uptake and toxic mechanism between species. For this purpose, the authors examined the

<p>correlations of toxicity to eight aquatic organisms: Theoretical considerations. <i>Science of the Total Environment</i> 408(20): 4549-4555.</p>	<p>interspecies correlations of toxicity between species of <i>Vibrio fischeri</i>, river bacteria, algae, <i>Daphnia magna</i>, carp, <i>Tetrahymena pyriformis</i>, fathead minnow and guppy based on the theoretical background. The results show that there are good interspecies correlations between marine bacterium and fresh water bacteria or fish and fish, which suggest that compounds share the same bio-uptake and toxic mechanism of action between the species. On the other hand, poor interspecies relationships were found between toxicities to algae and <i>T. pyriformis</i> or <i>D. magna</i>, which suggests that compounds have different toxic mechanisms of action between these species.</p> <p>The authors also show that interspecies relationships can be improved by inclusion of the octanol/water partition coefficient or the energy of the lowest unoccupied molecular orbital, which reflect the difference of bio-uptake or toxic mechanism of action between species for organic compounds. Benzoic acids show very different toxicity contributions to the three species, <i>V. fischeri</i>, <i>D. magna</i> and carp, which can be explain by the fact that they can be easily absorbed into the unicellular bacteria, <i>V. fischeri</i>, whereas the skin and lipid content of multicellular organisms (<i>D. magna</i> and fish) can strongly inhibit the bio-uptake for ionisable compounds. Good correlation coefficients were observed between toxicities to <i>V. fischeri</i> and <i>D. magna</i> or fishes by inclusion of hydrophobic and ionisation parameters. Thus, according to the authors, <i>V. fischeri</i> or <i>D. magna</i> can serve as a surrogate of fish toxicity for the hydrophobic and ionisable compounds studied. The authors also discuss toxic mechanisms of action based on the theoretical background of the interspecies correlation.</p>
<p>Cassani, S., S. Kovarich, E. Papa, P. P. Roy, L. van der Wal and P. Gramatica (2013). <i>Daphnia</i> and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity-activity modelling. <i>Journal of Hazardous Materials</i> 258-259: 50-60.</p>	<p><i>See above - QSAR</i></p>
<p>Raimondo, S., D. N. Vivian and M. G. Barron (2010). Web-based Interspecies Correlation Estimation (Web-Ice) for acute toxicity: User manual. Version 1.1. EPA/600/R-10/004 Gulf Breeze, F.L. http://www.epa.gov/ceampubl/fchain/webice/index.html</p>	<p><i>See above - QSAR</i></p>
<p>Weltje, L., P. Simpson, M. Gross, M. Crane and J. R. Wheeler (2013). Comparative acute and chronic sensitivity of fish and amphibians: a critical review of data. <i>Environmental Toxicology and Chemistry</i> 32(5): 984-994.</p>	<p>The objective of this study was to compare the relative sensitivity of amphibians and fish to chemicals. Acute and chronic toxicity data were obtained from the U.S. Environmental Protection Agency (U.S. EPA) ECOTOX database and from the scientific and regulatory literature. The overall outcome is that fish and amphibian toxicity data are highly correlated and that fish are more sensitive (both acute and chronic) than amphibians. In terms of acute sensitivity, amphibians were between 10- and 100-fold more sensitive than fish for only four of 55 chemicals and more than 100-fold more sensitive for only two chemicals. Concerning</p>

	<p>chronic toxicity data, amphibians were between 10- and 100-fold more sensitive than fish for only two substances (carbaryl and dexamethasone) and greater than 100-fold more sensitive for only a single chemical (sodium perchlorate). However, the comparison for carbaryl was subsequently determined to be unreliable and that for sodium perchlorate is a potential artefact of the exposure medium. Only a substance such as dexamethasone, which interferes with a specific aspect of amphibian metamorphosis, might not be detected using fish tests. However, several other compounds known to influence amphibian metamorphosis were included in the analysis, and these did not affect amphibians disproportionately. According to the authors, the results suggest that additional amphibian testing is not necessary during chemical risk assessment.</p>
<p><i>Threshold of toxicological concern</i></p>	
<p>de Wolf, W., A. Siebel-Sauer, A. Lecloux, V. Koch, M. Holt, T. Feijtel, M. Comber and G. Boeijs (2005). Mode of action and aquatic exposure thresholds of no concern. <i>Environmental Toxicology and Chemistry</i> 24(2): 479-485.</p>	<p>Thresholds of toxicological concern are based on the possibility of establishing an exposure threshold value for chemicals below which no significant risk is to be expected. The authors addressed environmental thresholds of no toxicological concern for freshwater systems (ETNCaq) for organic chemicals. They analysed several environmental toxicological databases (e.g. ECETOC, EURATS, US EPA fathead minnow, University of Utrecht) for acute and chronic endpoints and substance hazard assessments. Lowest numbers and 95th-percentile values were derived using data stratification based on mode of action (MOA) according to Verhaar (1 - inert chemicals; 2 - less inert chemicals; 3 – reactive chemicals; 4 - specifically acting chemicals). The ETNCaq values were derived by multiplying the lowest 95th percentile values with appropriate application factors; ETNCaq for MOA1–3 is approximately 0.1 µg/L. A preliminary analysis with complete MOA stratification of the databases shows that in the case of MOA1 or MOA2, the ETNCaq value could be even higher than 0.1 µg/L. For MOA4, a significantly lower ETNCaq value was observed based on the long-term toxicity information in the ECETOC database. The authors propose that the application of the ETNCaq value in a tiered risk-assessment scheme may help chemical producers to set data-generation priorities and to refine or reduce animal use. Further, it may help to inform downstream users concerning the relative risk associated with their specific uses and be of value in putting environmental monitoring data into a risk-assessment perspective.</p>
<p>Gross, M., K. Daginnus, G. Deviller, W. de Wolf, S. Dungey, C. Galli, A. Gourmelon, M. Jacobs, P. Matthiessen, C. Micheletti, E. Nestmann, M. Pavan, A. Paya-Perez, H.-T. Ratte, B. Safford, B. Sokull-Kluttgen, F. Stock, H.-C. Stolzenberg, J. Wheeler, M. Willuhn, A. Worth, J. M. Z. Comenges and M. Crane (2010).</p>	<p>This paper summarises the outcome of a workshop of regulatory, industry and academic scientists held to discuss the use of the threshold of toxicological concern (TTC) concept in aquatic environmental risk assessment, and in particular for endocrine active substances (EAS). A case study examining the use of the TTC for an EAS with an estrogenic MOA formed the basis for the discussions on the feasibility and acceptability, general advantages</p>

<p>Thresholds of toxicological concern for endocrine active substances in the aquatic environment. Integrated environmental assessment and management 6(1): 2-11.</p>	<p>and disadvantages, and the specific issues that need to be considered when applying the TTC concept for EAS in risk assessment. Issues surrounding the statistical approaches used to derive TTCs were also discussed. The participants concluded that the reliable use of a TTC in environmental risk assessment will require understanding of an untested substance's MOA and potency class, plus a reliable training set of data for long-term demographic effects on sensitive organisms of substances with that specific MOA, and that these are significant limiting factors for the applicability of this concept for EAS at the present time.</p> <p>To investigate further the applicability of the TTC concept for EAS, and to help develop the basis for eventual guidance, they recommended several follow-up activities: 1) further investigation into the combined use of <i>in vitro</i> and structure-activity data in establishing MOAs; 2) establishing criteria for the type, number, and acceptability of data when deriving a TTC, and for its subsequent use with a "new" substance with a reliably identified MOA; 3) establishing criteria for estimating reliable exposure concentrations of that substance for comparison with the TTC; and 4) investigation of whether an approach analogous to the use of toxic equivalents could be incorporated into use of the TTC for substances with a similar MOA to address mixture toxicity issues.</p>
<p>Tolls, J., M. Muller, A. Willing and J. Steber (2009). A new concept for the environmental risk assessment of poorly water soluble compounds and its application to consumer products. Integrated environmental assessment and management 5(3): 374-378.</p>	<p>The authors propose the use of the aquatic exposure threshold of no concern (ETNCaq; i.e., a concentration below which no adverse effects on the environment are to be expected) as described by de Wolf et al. (2005) for the environmental risk assessment of consumer products and in particular their lipophilic, poorly soluble ingredients representing large-volume substances whose aquatic toxicity cannot be adequately determined with standard methods for a number of reasons.</p> <p>For this purpose, the ETNCaq value of these poorly soluble substances is compared with the aquatic exposure levels. Aquatic exposure levels of substances with water solubility below the ETNCaq will not exceed the ecotoxicological no-effect concentration; therefore, their risk can be assessed as being negligible. The ETNCaq value relevant for substances with a narcotic mode of action is 1.9 µg/L. To apply the above risk assessment strategy, the solubility in water needs to be known. Most frequently, this parameter is estimated by means of quantitative structure/activity relationships based on the log octanol-water partition coefficient (log Kow). The predictive value of several calculation models for water solubility has been investigated by this method with the use of more recent experimental solubility data for lipophilic compounds. A linear regression model was shown to be the most suitable for providing correct predictions without underestimation of real water solubility. To define a log Kow threshold suitable for reliably predicting a water solubility of less than 1.9 µg/L, a confidence limit was established by statistical comparison of the experimental solubility data</p>

	<p>with their log Kow. It was found that a threshold of log Kow = 7 generally allows discrimination between substances with solubility greater than and less than 1.9 µg/L. Accordingly, organic substances with a baseline toxicity and log Kow > 7 do not require further testing to prove that they have low environmental risk. In applying this concept, the uncertainty of the prediction of water solubility can be accounted for. If the predicted solubility in water is to be below ETNCaq with a probability of 95%, the corresponding log Kow value is 8.</p>
<p>Mons, M. N., M. B. Heringa, J. van Genderen, L. M. Puijker, W. Brand, C. J. van Leeuwen, P. Stocks, J. P. van der Hoek and D. van der Kooij (2013). Use of the Threshold of Toxicological Concern (TTC) approach for deriving target values for drinking water contaminants. <i>Water Research</i> 47: 1666-1678.</p>	<p>Improving analytical techniques reveal more and more anthropogenic substances in drinking water and treated water. Most of the substances detected lack toxicity data to derive safe levels and have not yet been regulated, and although the concentrations found usually do not have adverse health effects, those substances are still undesired because of customer perception. This leads to the question how sensitive analytical methods need to become for water quality screening, at what levels water suppliers need to take action and how effective treatment methods need to be designed to remove contaminants sufficiently. In the Netherlands a clear and consistent approach called 'Drinking Water Quality for the 21st century (Q21)' has been developed. Target values for anthropogenic drinking water contaminants were derived by using the recently introduced Threshold of Toxicological Concern (TTC) approach. The target values for individual genotoxic and steroid endocrine chemicals were set at 0.01 mg/L. For all other organic chemicals the target values were set at 0.1 mg/L. The target value for the total sum of genotoxic chemicals, the total sum of steroid hormones and the total sum of all other organic compounds were set at 0.01, 0.01 and 1.0 mg/L, respectively. The Dutch Q21 approach is further supplemented by the standstill-principle and effect-directed testing. The approach is helpful in defining the goals and limits of future treatment process designs and of analytical methods to further improve and ensure the quality of drinking water, without going to unnecessary extents.</p>
<p><i>Integrated testing strategies</i></p>	
<p>ECETOC (2005). Workshop on Alternative Testing Approaches in Environmental Risk Assessment. . ECETOC Workshop Report No 5: 34 pp.</p>	<p>This report summarises possibilities to reduce the number of fish in environmental risk assessment, i.e. acute and chronic fish toxicity testing, fish bioconcentration/bioaccumulation as discussed at an ECETOC workshop in 2004. The status of testing methods under development (fish cell lines, fish embryos, omics, reduction strategies as the threshold/step-down approach, review of guidelines for chronic fish toxicity testing) was discussed as well as the use of QSARs for acute fish toxicity testing, and recommendations for further research (e.g. call for CEFIC LRi ECO 8 project drafted at this workshop) were given. Regarding bioconcentration/bioaccumulation the focus was on tiered testing strategies and improvement of <i>in silico</i> prediction models with ADME information. A lot of the recommendations had</p>

<p>ECETOC (2007). Intelligent testing strategies in ecotoxicology: mode of action approach for specifically acting chemicals. Technical Report No. 102: 182 pp.</p>	<p>been followed up with several activities and/or introduced into the regulatory frameworks.</p> <p>This report provides an in-depth assessment of the current science underlining the use of mode of action (MOA) information for specifically acting chemicals. The MOA describes the understanding of selected key events that leads to toxic effects. After a review of different MOA classification schemes, the Verhaar et al. (1992) approach, based on acute effect information and comprising 4 class of chemicals according to their MOA (MOA1 to MOA4), was adopted as a starting point. To date, the MOA4 (specifically acting chemicals with selective biomolecular interactions) has not been utilized for chronic assessment. Considering both acute and chronic effects, the taskforce has therefore extended the MOA4 class into 4 pharmacological sub-classes based on: (a) protein receptors, (b) enzymes, (c) ions channels and (d) transporters, which provide a consistent a consistent basis to review the MOA for both therapeutic (intended) effects and for toxicity (e.g. MOAs for neurotoxicity and tumour induction by some carbamates).</p> <p>Expanding upon published intelligent testing strategy (ITS) schemes the Task Force has developed a simple five step flow-chart which starts with data gathering on physical-chemical data, SAR predictions, <i>in vitro</i> tests, and data from mammalian studies for the compound of interest and for related chemicals. Evaluation of these data provides information on exposure to, and possible MOA of, the substance. Data may come from efficacy or therapeutic data, or from read-across within chemical classes. Concerning effects, valuable guidance may be obtained from molecular, biochemical or cellular toxic responses measured in both <i>in vivo</i> and <i>in vitro</i> studies.</p> <p>This culminates with a pragmatic, stepwise prioritisation for the assessment of chronic effects in regulatory aquatic test species, providing guidance on the selection from microbial, plant, invertebrate or vertebrates. The application of the 'specific MOA flow-chart' is illustrated through five case studies: (i) Ion channel mediated effect case study and cypermethrin, (ii) Receptor mediated effect case study and 17a-ethinylestradiol (EE2), (iii) Transporter protein mediated effects case study and fluoxetine, (iv) Enzyme mediated effect case study and ketoconazole, (v) Microbial enzyme inhibitor case study and triclosan.</p> <p>As a conclusion, the Task Force highlight the fact that depending on the exposure conditions, some chemicals may induce biological effects that suggest more than one MOA (e.g. carbamates), thus caution is needed when using the protein target ITS approach to guide test protocol design. Therefore, the Task Force recommends an ITS approach to aquatic ecotoxicity testing that includes the following key elements:</p> <ul style="list-style-type: none"> - Gather MOA information on the primary pharmacological/toxicological activity for the chemical of interest for the target species as well as mammalian data, also considering
---	---

structurally-related chemicals, to be used in a weight-of-evidence type of approach;

- Make use of non-traditional sources of biological information, especially the growing biomedical and 'omics' electronic databases on zebrafish, marine invertebrates and other non-mammalian species;
- If there is evidence for the main MOA being via a protein target, use this insight to guide the efficient selection of regulatory test methods;
- Measure biomarker responses (e.g. vitellogenin) if desired for read-across purposes or setting test concentrations, however, focus on population relevant endpoints (survival, development growth and reproduction) for generating NOEC (or EC_x) values or calculation of PNEC values for application in environmental risk assessment;
- Be cautious of using acute interspecies sensitivity ratios (ISRs) for algae, crustaceans or fish, since the available data suggests they are of limited value for ITS application, presumably since acute high levels exposure induces different MOAs compared to chronic low level exposures.

Finally, the Task Force has identified key knowledge gaps around regulatory test species which create major uncertainty in developing the ITS approach for many MOA4 chemicals. Therefore, the Task Force has also identified five research needs that will help reduce the scale of this problem in the ITS context. These recommendations are:

- To critically review data (especially chronic studies) from a wider range of chemicals in the context of the proposed MOA and ITS framework, including chemicals where the mammalian MOA is less specific than for agrochemicals and pharmaceuticals;
- To develop aquatic plant ADME models with special regard to understanding key biotransformation enzymes;
- To strengthen the use of small invertebrate models by investing in a hierarchy of biological understanding (including genomics, proteomics and population responses) in the commonly used freshwater and marine invertebrate species (including both arthropods and non-arthropods);
- For animal welfare reasons, to minimise the need for *in vivo* fish bioconcentration testing by developing *in vitro* fish protocols for chemical metabolism and by developing a small-scale invertebrate bioconcentration test method;
- To support risk assessments of endocrine disrupters, develop a database of the normal (baseline) range for developmental and reproductive endpoints in aquatic organisms measured across different laboratories;
- To capitalise on the learning from zebrafish biomedical and 'omic' research, there would be value in the establishment of a publicly available database on zebrafish ADME and toxicity

	information; -To support continuing research into SAR validation by providing 'training data sets' based on high quality <i>in vivo</i> chronic tests with plants, invertebrates and fish.
Grindon, C., R. Combes, M. T. D. Cronin, D. W. Roberts and J. F. Garrod (2008). Integrated decision-tree testing strategies for environmental toxicity with respect to the requirements of the EU REACH legislation. <i>Alternatives to Laboratory Animals</i> 36 Suppl 1: 29-42.	In the light of the, at that time, upcoming REACH legislation, the authors discuss possibilities to reduce the number of fish for environmental toxicity testing, evaluate the status of alternative methods and present integrated decision-tree testing strategies covering acute and chronic fish toxicity as well as bioconcentration/bioaccumulation. The testing strategy is based on the use of available information on the chemical (including mammalian toxicity), exposure, <i>in silico</i> , <i>in vitro</i> , fish embryos, bacterial based assays, threshold/step-down approach for acute toxicity testing. Before carrying out a chronic fish test the authors recommend estimating the bioaccumulation (BCF models or fish bioconcentration test). Only if long-term exposure and bioaccumulation may occur, chronic fish test are recommended by the authors.
OECD (2012a). Fish toxicity testing framework. Series on Testing and Assessment No. 171; ENV/JM/MONO(2012)16; OECD, Paris, France.	This document presents a review of fish toxicity testing for the regulatory purpose of chemical safety as proposed by the USA in 2008. The main focus is on fish toxicity with an emphasis on endocrine disruptors, but fish bioaccumulation is also considered where relevant. The document was initially elaborated by a group of experts and reviewed at an OECD Workshop on a Fish Toxicity Testing Framework, held on 28-30 September 2010 (UK). A review of regulatory needs for fish tests under various jurisdictions in OECD countries is provided, followed by a review of statistical issues and general test considerations. The document examines animal welfare concerns and alternatives and provides a systematic review of existing and draft OECD Guidelines which use fish for toxicity or bioaccumulation studies. A generic framework for assessing the environmental hazards of chemicals using fish tests in the most efficient way. An Annex contains conclusions and recommendations made and agreed at the workshop in September 2010. The recommendations concern, among other aspects, possible improvements to existing Test Guidelines, development of guidance on specific issues, harmonisation of existing Test Guidelines for common issues, development of new Test Guidelines, and proposals for deletion of outdated Test Guidelines. Recommendations resulted in the deletion of OECD TG 204 (see above) and revision of OECD TG 210.
Scholz, S., E. Sela, L. Blaha, et al. (2013). A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. <i>Regulatory Toxicology and Pharmacology</i> 67(3): 506-530.	The authors provide an overview on current regulatory requirements for animal tests in environmental hazard and risk assessment of chemicals, plant protection products, pharmaceuticals, biocides, feed additives and effluents. They discuss replacement, reduction and refinement of animal tests (e.g using fish, birds, amphibian) covering acute and chronic toxicity (incl endocrine disruptors), and bioaccumulation. Perspectives and limitations of

	<p>alternative approaches are discussed. Free access to existing (proprietary) animal test data, availability of validated alternative methods and a practical implementation of conceptual approaches such as the Adverse Outcome Pathways and Integrated Testing Strategies were identified as major requirements towards the successful development and implementation of alternative approaches. Although this article focusses on European regulations, its considerations and conclusions are of global relevance.</p>
<p><i>In vitro methods for fish bioaccumulation testing – improvement of prediction models</i></p>	<p>This topic is not further discussed in the report, however, it is an important topic and industry is very much interested in <i>in vitro</i> biotransformation assays, e.g. use of fish hepatocytes or cellular subfractions as the fish S9 (see below Johannning et al., 2012). A lot of efforts have been put into the development of <i>in vitro</i> to <i>in vivo</i> extrapolation models and refined bioconcentration/bioaccumulation prediction models (see recent publication of Nichols et al., 2013).</p>
<p>Johanning, K., G. Hancock, B. Escher, A. Adekola, M. J. Bernhard, C. Cowan-Ellsberry, J. Domoradzki, S. Dyer, C. Eickhoff, M. Embry, S. Erhardt, P. Fitzsimmons, M. Halder, J. Hill, D. Holden, R. Johnson, S. Rutishauser, H. Segner, I. Schultz and J. Nichols (2012). Assessment of metabolic stability using the rainbow trout (<i>Oncorhynchus mykiss</i>) liver S9 fraction. Current Protocols in Toxicology Suppl. 53: Chapter: 14.10.11-14.10.28, August 2012.</p>	<p>To summarise, standard protocols are given for assessing metabolic stability in rainbow trout using the liver S9 fraction, they address: the isolation of S9 fractions from trout livers, evaluation of metabolic stability using a substrate depletion approach, and expression of the result as <i>in vivo</i> intrinsic clearance. Additional guidance is provided on the care and handling of test animals, design and interpretation of preliminary studies, and development of analytical methods. Although initially developed to predict metabolism impacts on chemical accumulation by fish, these procedures can be used to support a broad range of scientific and risk assessment activities including evaluation of emerging chemical contaminants and improved interpretation of toxicity testing results. These protocols have been designed for rainbow trout and can be adapted to other species as long as species-specific considerations are modified accordingly (e.g., fish maintenance and incubation mixture temperature). Rainbow trout is a cold-water species. Protocols for other species (e.g., carp, a warm-water species) can be developed based on these procedures as long as the specific considerations are taken into account.</p> <p>The detailed protocols (including the calculation tool) are available from: http://onlinelibrary.wiley.com/doi/10.1002/0471140856.tx1410s53/full</p>
<p>Nichols J.W., Huggett D.B., Arnot J.A., Fitzsimmons P.N., Cowan-Ellsberry C.E. (2013). Toward improved models for predicting bioconcentration of well-metabolized compounds by rainbow trout using measured rates of <i>in vitro</i> intrinsic clearance. Environ Toxicol Chem. 32(7):1611-22.</p>	<p>The authors developed to predict the bioconcentration of well-metabolized chemicals by rainbow trout. The models employ intrinsic clearance data from <i>in vitro</i> studies with liver S9 fractions or isolated hepatocytes to estimate a liver clearance rate, which is extrapolated to a whole-body biotransformation rate constant (kMET). Estimated kMET values are then used as inputs to a mass-balance bioconcentration prediction model. An updated algorithm based on measured binding values in trout is used to predict unbound chemical fractions in blood, while other model parameters are designed to be representative of small fish typically used in</p>

whole-animal bioconcentration testing efforts. Overall model behaviour was shown to be strongly dependent on the relative hydrophobicity of the test compound and assumed rate of *in vitro* activity. The results of a restricted sensitivity analysis highlight critical research needs and provide guidance on the use of *in vitro* biotransformation data in a tiered approach to bioaccumulation assessment.

The authors provide supplemental data on estimation of S9 protein content and scaling of specific metabolic rates among fish of different sizes.

Appendix 1 Weblinks to general information on non-standard methods

This compilation is not intended to be complete but it may serve as a starting point for further information in the field of alternatives to animal testing.

Web address	Brief description
http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam	EURL ECVAM website
http://ecvam-dbalm.jrc.ec.europa.eu	EURL ECVAM DataBase service on ALternative Methods to animal experimentation (DB-ALM)
http://bookshop.europa.eu/en/the-eurl-ecvam-search-guide-pbLBN124391	The DB-ALM search guide
http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/laboratories-research/predictive_toxicology	JRC Predictive Toxicology
http://www.oecd.org/env/testguidelines	OECD Test Guidelines Programme
http://ec.europa.eu/dgs/health_consumer/index_en.htm	DG SANCO
http://ec.europa.eu/dgs/environment/index_en.htm	DG ENV
http://ec.europa.eu/enterprise/index_en.htm	DG ENTR
http://echa.europa.eu	ECHA
http://www.efsa.europa.eu	EFSA
http://www.jacvam.jp/en/index.html	Japanese Center for the Validation of Alternative Methods (JaCVAM)
http://iccvam.niehs.nih.gov	The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) (USA)
http://www.niehs.nih.gov/research/atniehs/dntp/assoc/niceatm	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)
http://www.seurat-1.eu	Seurat-1
http://www.cosmostox.eu/home/welcome	COSMOS
http://www.detect-iv-e.eu	DETECTIVE
http://www.alttox.org	ALTTOX
http://www.journals.elsevier.com/toxicology-in-vitro	Toxicology in Vitro
http://altweb.jhsph.edu/pubs/journals/atla/issue.html	ATLA: Alternatives to Lab Animals
http://www.altex.ch/Home.12.html	ALTEX: Alternatives to Animal Experimentation
http://www.estiv.org	The European Society of Toxicology In Vitro
http://www.icare-worldwide.org	International Centre for Alternatives in Research and Education
https://www.cosmeticseurope.eu/index.php	Cosmetics Europe (CE)
http://ec.europa.eu/enterprise/epaa	The European Partnership for Alternative Approaches to Animal Testing (EPAA)
http://www.bfr.bund.de/en/zebet-58194.html	BfR, ZEBET
http://www.nc3rs.org.uk	National Centre for the Replacement, Refinement and Reduction of Animals

Web address	Brief description
	in Research
http://www.forschung3r.ch	3R Research Foundation Switzerland
http://toxnet.nlm.nih.gov/altbib.html	Resources for Alternatives to the Use of Live Vertebrates in Biomedical Research and Testing
http://www.gopubmed.org/web/go3r	Transinsight GmbH
http://www.epa.gov/ncct/Tox21	Tox21

Appendix 2 Status of test method submissions to EURL ECVAM

Table updated as of 31/07/2014

TM no	Health effect or environmental effect or other	TM description	Assessment of pre-submission	Outcome of pre-submission assessment	Assessment status of full submission	Outcome of full submission assessment*	Validation and acceptance status
TM2014-03	Respiratory permeability/penetration, toxicokinetics	In vitro 3D reconstructed human airway epithelial model for evaluation of substances through the respiratory track	Ongoing				
TM2014-02	Genotoxicity	gHistone 2AX coupled with in cell western technique, as biomarker of genotoxicity	On hold	Request of additional information			
TM2014-01	Fish acute toxicity	Fish cell line acute toxicity test using Rainbow Trout (RT) gill cell line and using a combination of fluorescent dyes in order to measure cell viability and establish EC50 value	Finalised	Invitation to send a full submission			
TM2013-03	Skin sensitisation	Test method for the assessment of the skin sensitization potential of chemicals based on transcriptomics and intracellular signal transduction analysis in a mouse dendritic cell-like cell line	Finalised	Not considered for follow-up			
TM2013-02	Skin sensitisation	Updated Myeloid U937 Skin Sensitisation Test	Finalised	Invitation to send a full submission			

TM2013-01	Intestinal permeability/penetration, toxicokinetics	In vitro assay to measure the permeability of compounds (expressed as apparent permeability- Papp value) through the human intestinal Caco-2 cell line	Finalised	Not considered for follow-up			
TM2012-06	Carcinogenicity	CTA test method using Bhas 42 cell line for carcinogenicity assessment, based on a two-component protocol: (i) initiation assay at low cell density; (ii) promotion assay at higher cell density.	n.a.		n.a.		Validation finalised, ECVAM Recommendation published
TM2012-05	Skin sensitisation	Method for assessing the skin sensitization potential and potency of chemicals based on IL-18 release and cytotoxicity measurements in reconstituted human epidermis (RHE)	Finalised	Invitation to send a full submission			
TM2012-04	Reproductive toxicity - Embryotoxicity	Stem cell Test method for the identification of embryotoxic hazards for humans on the basis of the analysis of molecular end-points in D3 mouse embryonic stem cells under differentiation	Finalised	Not considered for follow-up			
TM2012-03	Skin sensitisation	Test method for the assessment of the skin sensitization potential of substances based on the assessment of histological damage in	Finalised	Not considered for follow-up			

		human skin explants					
TM2012-02	Skin permeability/penetration, toxicokinetics	In vitro permeability method employing an artificial skin-mimetic membrane to predict skin penetration of active pharmaceutical ingredient solutions or test formulations	On hold	Context of test method to be clarified			
TM2012-01	Quality control of vaccines	Combined assay for detection of tetanus toxicity in tetanus vaccines	Finalised	Invitation to send a full submission			
TM2011-14	Sensitisation	In vitro testing strategy based on the combination of information on peptide reactivity, responses in keratinocytes and activation of Dendritic cells for the identification of skin sensitisation hazard	Finalised	Not considered for follow-up, skin sensitisation IATAs will be addressed in the OECD Task Force on Hazard Assessment			
TM2011-13	Eye irritation	Extracellular acidification test for water-soluble severe or non-classified eye irritants	On hold	Request of additional information			

TM2011-12	Sensitisation	Two-tiered approach to determine the skin sensitising capacity and potency of chemicals	Finalised	Not considered for follow-up because an improved version of the test method has been submitted under TM2012-05			
TM2011-11	Skin sensitisation	In vitro method based on gene expression in 3D reconstructed epidermis for identifying skin sensitisers	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Under external validation
TM2011-10	Skin sensitisation	In vitro test measuring luciferase induction in a human cell line for the testing of skin sensitisers	Finalised	Invitation to send a full submission			
TM2011-09	Sensitisation	In vitro test for the detection of sensitisers based on the analysis of the relative expression levels of a biomarker signature of 200 genes using microarrays	Finalised	Invitation to send a full submission	Finalised	Does not qualify for entering validation	
TM2011-08	Sensitisation	In vitro test for the detection of sensitisers based on the quantification of protein biomarkers in Dendritic Cells models using mass spectrometric assays	On hold	Request of additional information			
TM2011-07	Sensitisation	In vitro method for the identification of sensitisers based on the evaluation of Dendritic Cells migration in a two-	Finalised	Not considered for follow-up			

		compartment system					
TM2011-06	Eye irritation	In vitro prediction of acute ocular irritation	Finalised	Invitation to send a full submission			
TM2011-05	Skin irritation	In vitro Skin Irritation Test with an open source reconstructed epidermis	Finalised	Invitation to send a full submission	Ongoing for a revised full submission		
TM2011-04	(Not specified)	Bioaccessibility testing from powder and massive samples of metals and alloys	Finalised	Not considered for follow-up (no regulatory relevance with regard to impact on the 3Rs)			
TM2011-03	Reproductive toxicity - Endocrine disruption	Yeast Estrogen Screening Assay (YES) and Yeast Androgen Screening Assay (YAS)	Finalised	Invitation to send a full submission	Finalised for YAS	Request of additional information	
TM2011-02	Skin sensitisation	In vitro Method for Identifying Skin Sensitisers Combining Peptide Binding with ARE/EpRE Mediated Gene Expression in Human Skin Cells	Finalised	Invitation to send a full submission			
TM2011-01	Acute toxicity, chronic toxicity, genotoxicity	Dequenching after photobleaching cytotoxicity test	On hold	Request of additional information			
TM2010-08	Genotoxicity	Transcriptomics-based Genomics-Genotox assay - Gene expression profiling in HepG2 cells for in vivo genotoxicity prediction	Finalised	Invitation to send a full submission	Finalised	Does not qualify for entering validation	

TM2010-07	Reproductive Toxicity - Endocrine disruption	Transcriptional activation assay for detection of (anti-)androgenic activity of chemicals	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Under validation
TM2010-06	Neurotoxicity	Embryonic Rat Dorsal Root Ganglia Organotypic Culture	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Not prioritised for validation by ECVAM
TM2010-05	Reproductive toxicity - Meiosis	Bovine Oocyte Maturation Assay	Finalised	Invitation to send a full submission			
TM2010-04	Eye irritation	In vitro method to detect changes in the metabolic rate of cells	On hold	Request of additional information			
TM2010-03	Skin sensitisation	In vitro test measuring luciferase induction in a human cell line for the testing of skin sensitisers	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering peer review	Validation finalised, ECVAM Recommendation published
TM2010-02	Reproductive toxicity - Endocrine disruption	Transactivation assay for detection of compounds with (anti)androgenic potential using PALM cells	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Not prioritised for validation by ECVAM due to issues regarding the material transfer agreement
TM2010-01	Reproductive toxicity - Endocrine disruption	Rat recombinant androgen receptor binding assay for the detection of compounds with (anti)androgenic potential	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Not prioritised for validation by ECVAM
TM2009-12	Neurotoxicity	In vitro approach for organophosphorous compounds-induced neurotoxicity	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Not prioritised for validation by ECVAM

TM2009-11	Reproductive toxicity - Endocrine disruption	Transcriptional activation assay for detection of (anti-)estrogenic activity of chemicals	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering performance standards based validation	Not prioritised for validation by ECVAM
TM2009-10	Skin irritation	Reconstructed human Epidermis test method for In vitro skin irritation testing	Finalised	Invitation to send a full submission			
TM2009-09	Skin irritation	Epidermal skin irritation test	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering peer review	
TM2009-08	Genotoxicity	Human lymphoblastoid TK6 cells transfected with the GADD45a-green fluorescent protein for detecting genome damage and genotoxic stress	Finalised	Invitation to send a full submission	Finalised	Consultation ongoing with the Preliminary Assessment of Regulatory Relevance (PARERE) network	
TM2009-07	Cardiotoxicity	Automated screening test to detect cardiotoxicity in zebrafish	On hold	Request of additional information			
TM2009-06	Skin sensitisation	Direct Peptide Reactivity Assay for Screening the Skin Sensitisation Potential of Chemicals	n.a.		Finalised	Qualifies for entering validation	Validation finalised, ECVAM Recommendation published
TM2009-05	Skin sensitisation	Myeloid U937 Skin Sensitisation Test	n.a.		Finalised	Qualifies for entering validation	Validation stopped
TM2009-04	Eye irritation	Chemical reactivity measurement using the glutathione (GSH and GSSG) peptide binding HPLC assay	n.a.		Finalised	Qualifies for entering validation	Not prioritised for validation by ECVAM

TM2009-03	Eye irritation	Chemical reactivity measurement using the cysteine or lysine peptide binding HPLC assay	n.a.		Finalised	Qualifies for entering validation	Validation stopped
TM2009-02	Reproductive toxicity - Endocrine disruption	Transactivation assay for detection of compounds with (anti)estrogenic potential using MELN cells	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Under validation
TM2009-01	Eye irritation	In vitro assay to predict the ocular irritation potential by measuring protein denaturation	Finalised	Invitation to send a full submission	Ongoing		
TM2008-13	Skin absorption/penetration	In vitro diffusion method for measuring skin absorption	Finalised	Invitation to send a full submission			
TM2008-12	Acute and chronic toxicity	Quantitative live cell cytotoxicity assay based on measurements of DNA alterations (later resubmitted as DAP)	On hold	Request of additional information			
TM2008-11	Skin sensitisation	Test method for assessing the sensitization potential of proteins associated with type I (IgE mediated) hypersensitivity reactions	Finalised	Not considered for follow-up			
TM2008-10	Eye irritation	Test method based on reconstructed human tissue model	n.a.		Finalised	Qualifies for entering validation	Validation finalised, under ESAC peer review
TM2008-09	Eye irritation	Test method based on human corneal epithelium	n.a.		Finalised	Qualifies for entering validation	Validation stopped
TM2008-08	Genotoxicity	Assay for the assessment of DNA-modifying agents	Finalised	Not considered for follow-up			

TM2008-07	Reproductive toxicity - Endocrine disruption	Test for edocrine disruptor identification	Finalised	Invitation to send a full submission			
TM2008-06	Eye irritation	Chorioallantoic membrane vascular assay	Finalised	Invitation to send a full submission			Under external validation
TM2008-05	Skin sensitisation	Human Cell Line Activation Test	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Validation finalised, ESAC peer review finalised Draft EURL ECVAM recommendation under restricted commenting
TM2008-04	Skin sensitisation	Test method based on gene-expression analysis in human Dendritic Cells for discriminating sensitisers from non sensitisers	Finalised	Invitation to send a full submission	Finalised	Qualifies for entering validation	Under external pre-validation
TM2008-03	Eye irritation	Porcine corneal opacity reversibility assay	Finalised	Invitation to send a full submission			
TM2008-02	Skin irritation	Test method based on reconstructed human epidermis	n.a.		Finalised	Qualifies for entering validation	Validation finalised, OECD Test Guideline / EU Test Method available
TM2008-01	Skin irritation		n.a.		Finalised	Qualifies for entering validation	Validation finalised, OECD Test Guideline / EU Test Method available

*Outcome of full submission assessment:

"Qualifies for entering validation" or "Qualifies for entering performance standards-based validation" or "Qualifies for entering peer review" means that the test method is ready to enter a (performance standards-based) validation or peer review process. It does, however, not mean that a test method, which qualifies for entering validation, will automatically be validated by EURL ECVAM.

EURL ECVAM regularly prioritises all test methods which have been submitted and found to qualify for entering validation, against pre-defined criteria, and decides which test methods/approaches will finally enter validation by EURL ECVAM in the light of available capacities. It is foreseeable that not all submitted test methods can be validated.

In some cases, the test methods listed in this table may already undergo a validation study (by ECVAM or other organisations).

Appendix 3 Status of computational models in the JRC QSAR database

Documenting the validity of QSAR models

Although thousands of QSAR models have been developed and published in the scientific literature, and some models have been used in regulatory assessment of chemicals in some countries for many years, a transparent validation process and objective determination of the reliability of QSAR models are crucial to further enhance their regulatory acceptance.

In November 2004, the OECD Member Countries agreed on the principles for validating QSAR models for their use in the regulatory assessment of chemical safety. The internationally agreed principles provide Member Countries with a consistent and scientifically motivated framework for evaluating the regulatory applicability of QSAR models.

In February 2007, the OECD published a "Guidance Document on the Validation of (Q)SAR Models" with the aim of providing guidance on how specific QSAR models can be evaluated with respect to the OECD principles. A check list for the validation, a reporting format for the validation, and case studies are attached as annexes:

The OECD Principles for QSAR Model Validation, which are intended to be read in conjunction with the guidance document, are as follows:

"To facilitate the consideration of a (Q)SAR model for regulatory purposes, it should be associated with the following information:

- 1) a defined endpoint
- 2) an unambiguous algorithm
- 3) a defined domain of applicability
- 4) appropriate measures of goodness-of-fit, robustness and predictivity
- 5) a mechanistic interpretation, if possible"

The **QSAR Model Reporting Format (QMRF)** was developed by the JRC and EU Member State authorities as a harmonised template for summarising and reporting key information on QSAR models, including the results of any validation studies. The information is structured according to the OECD validation principles.

Introduction to the JRC QSAR Model Database

The **JRC QSAR Model Database** (<http://qsardb.jrc.ec.europa.eu/qmrf>) is a freely accessible web application that enables users to submit, publish, and search QMRF reports. Developers and users of QSAR models can submit to the dedicated mailbox information on QSARs by using the QMRF. A downloadable **QMRF editor** (<http://sourceforge.net/projects/qmrf/files/QMRF%20Editor/2.0.0>) is used for this purpose. The JRC then performs a quality control (i.e. adequacy and completeness of the documentation) of the QMRF submitted. Properly documented QMRFs are included in the JRC QSAR Model Database. Inclusion of the model does not imply acceptance or endorsement by the JRC or the European Commission, and responsibility for use of the models lies with the end-users.

Status of QMRFs in the JRC QSAR Model Database

At the time of writing (December 2013), the JRC QSAR Model Database contains 70 QMRFs (Figure 1). A number of additional QMRFs will also be uploaded in a new version of the database that will be available soon from the same webpage.

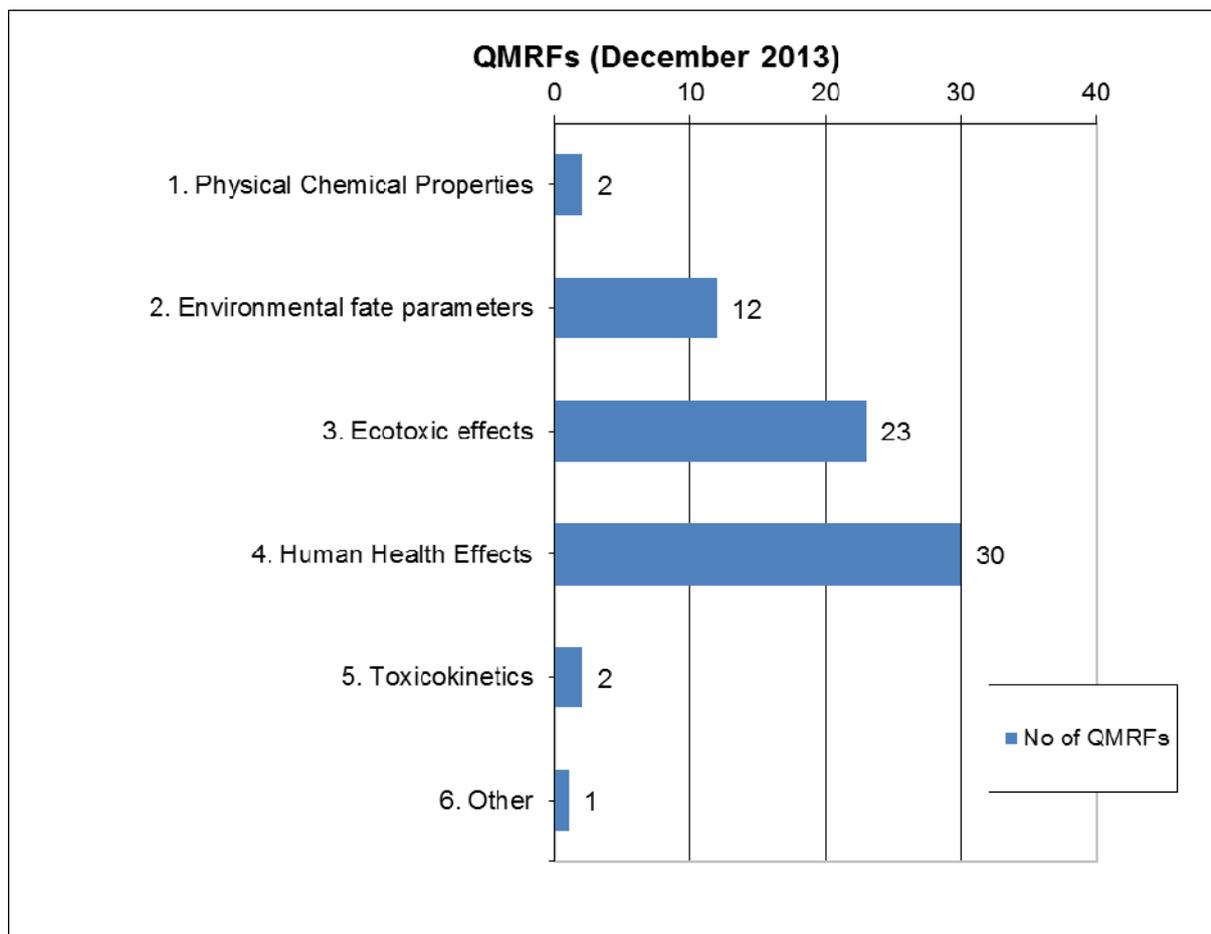


Figure 1. Status of QMRFs in the JRC QSAR Model Database

Reference

OECD (2007). Guidance Document No. 69 on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. ENV/JM/MONO(2007)2. Available from: <http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282007%292&do clanguage=en>

Appendix 4 Status of alternative methods in DB-ALM

Annett Janusch Roi

1. DB-ALM

Ready access to comprehensively and adequately described alternative methods is a prerequisite for their use within decision making processes by regulators and scientists or any end-user in the life sciences. The JRC has established and is managing the EU EURL ECVAM DataBase service on ALternative Methods to animal experimentation (DB-ALM), which provides standardised method descriptions and related information to enhance the knowledge about and the uptake of alternative and advanced methods at all stages of development and regulatory acceptance in the different policy areas and for different purposes.

DB-ALM originated from the Communication of the Commission to Council and European Parliament SEC (91)1794 and has further been reinforced by Directive 2010/63/EU.

1.1 Key Features

The DB-ALM is a freely accessible Internet based database service covering various aspects of alternative (principally, but not only, *in vitro*) methods, with focus on toxicity assessments of chemicals and/or formulations. The DB-ALM is accessible at:

<http://ecvam-dbalm.jrc.ec.europa.eu>

The key feature of the DB-ALM is to provide user-oriented documentation in the form of quality controlled descriptions of alternative approaches, prepared by experts and are consequently ready for immediate use (*factual* and *evaluated* information). The descriptions represent an outcome of extensive bibliographic reviews and/or direct contacts with the method developer/user.

The methods are described at various levels of detail and according to pre-determined criteria for data content to ensure consistent and adequately descriptions of methods. The respective reporting formats have evolved significantly since the first establishment of the DB-ALM in the early 2000's so as to capture all information elements necessary to allow judgments of its usefulness, that now cover information on the *potential of a method* including its intended objective and applications, its rationale and scientific principle, its relevance, a summary description of study results obtained so far including performance and reliability evaluations as available and appropriate, discussions on strengths and eventual limitations completed with their status of development, validation or regulatory acceptance.

These review documents are complemented with detailed step-by-step descriptions in the form of protocols, where available, to also allow the transfer of a technique to another laboratory.

1.2 Status

In terms of content, the DB-ALM currently provides the information summarised in Table 1:

Table 1. Information provided by DB-ALM

Information Sector	Data Sheet Number
Topic Summaries	5
Method Summaries	159
Protocols	150
Evaluation Studies	61
EU Integrated Project	2
Validation studies	19
Test Results (individual investigations)	9240
Persons & Institutions active in the field of alternative methods	242
Bibliographic References	6958

June 2014

In total, five topic areas in the form of thematic reviews are covered providing method summary descriptions, while individual protocols are made available for 26 topic areas addressing human health and eco-toxicological effects of chemical substances, mechanistic information, quality control of biological products, and biocompatibility and safety testing of medical devices (Table 2).

Over 150 biological endpoints are addressed referring to biological processes, responses or effects that can be measured at various levels of biological organisation.

The biological endpoints are both cytological and tissue-specific, including: a) interactions on the molecular level (including biochemistry and bio-kinetics); b) basal cytotoxicity testing; and c) functional parameters of organs and tissues.

1.3 Usage

Recent years have shown the consolidation of the DB-ALM where the online information content has been enhanced and/or revised by 33 % over the past four years leading to a constant average annual increase in new registrations to the service (Figure 1).

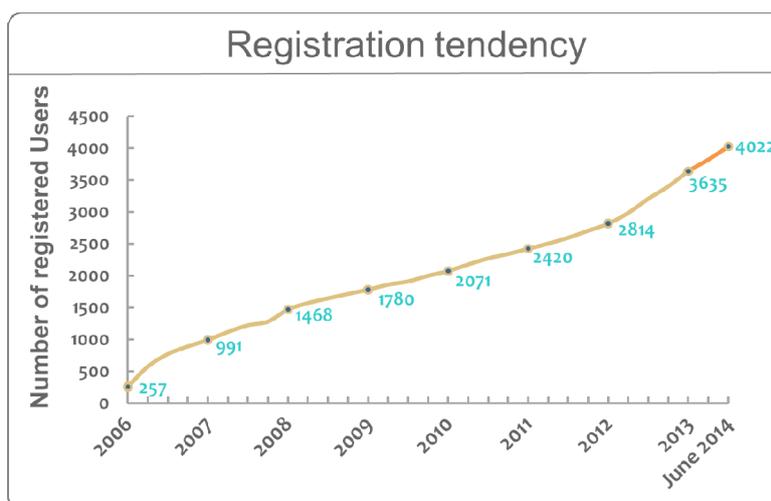


Figure 1. Usage of DB-ALM since 2006

In 2013 there was a doubling in the number of new registrations compared to the same period of the year before and there was a download of medially over 300 documents/months, representing an increase by 75% compared to the same period of 2012. In total, the DB-ALM can refer to over 4000 registered users from 82 countries covering representatives from academia (45%), industry (33%) and regulators (13%), the animal welfare movement and others (9 %).

In 2011, the AXLR8 coordination and support action (funded by DG Research & Innovation) that aims to accelerate the transition period versus a more sophisticated approach to chemical and product safety assessment, whilst reducing the animal use, indicated the DB-ALM as the platform where to store and disseminate their FP6 & FP7 research projects results regarding non-animal approaches.

In 2013 the OECD Advisory Group on Molecular Screening and Toxicogenomics has set up a drafting group to develop guidance for characterising non-guideline *in vitro* methods used for regulatory purposes as a supplement to the existing guidance for development and assessment of Adverse Outcome Pathways. The DB-ALM was considered as a potential resource for storing those methods in a standardised manner.

2. EURL ECVAM Search Guide

The EURL ECVAM Search Guide (SG) was developed to inform and support untrained database users to find high quality information on relevant alternative methods and strategies in the large amount of available information resources in an easy, yet systematic, and efficient way. This is most relevant where regulatory requirements mandate the application of the 3Rs, particularly, during the project preparations in biomedical sciences and toxicology. The SG is downloadable from: <http://bookshop.europa.eu>.

The SG provides search procedures, suggested search terms and user guidance to facilitate the location of the desired information on 3Rs animal alternatives in addition to an inventory of relevant information resources.

The SG was first published as a handbook in 2012 by the EU Bookshop where it was included at a certain point at 4th position in the Top 10 of the most downloaded publications of the EU Bookshop. Encouraged by this success, the JRC has re-published a second entirely updated version, publicly available since August 2013.

Table 2. DB-ALM Information Coverage – Method Summaries and Protocols⁵

Topic (field of application)	No. of Datasheets	
	Method Summary	Protocol
Acute Systemic Toxicity		6
Basal Cytotoxicity		17
Biocompatibility & Safety of Medical Devices		3
Biokinetics (Percutaneous Absorption)	11	3
Cancer Research (Drug Discovery and Activity Testing)	5	1
Carcinogenicity (Cell transformation and Tumour promotion)		5
Cardiotoxicity		2
Culture Methods		4
Digestive System Toxicity		1
Effects on Reproduction	38	14
Developmental toxicity	21	8
Effects on Female fertility	8	3
Effects on Male fertility	9	3
Endocrine Organs Toxicity		1
Environmental Toxicity (Aquatic Short-Term toxicity)		5
Genotoxicity/Mutagenicity		7
Haematotoxicity		1
Hepatotoxicity / Metabolism-mediated Toxicity	23	8
Immunotoxicity		1
Local Toxicity	71	53
Eye Irritation	70	32
Phototoxicity		9
Skin Irritation and Corrosivity	1	12
Myotoxicity		1
Nephrotoxicity		4
Neurotoxicity	7	1
Pyrogenicity testing		2
Respiratory Tract Toxicity)		1
Sensitisation and Allergy (Photoallergenicity, Skin sensitisation and allergic contact dermatitis)		3
Total	155	143

⁵ One or more protocols can belong to a method summary

Europe Direct is a service to help you find answers to your questions about the European Union
Freephone number (*): 00 800 6 7 8 9 10 11

(*): Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server <http://europa.eu>.

How to obtain EU publications

Our publications are available from EU Bookshop (<http://bookshop.europa.eu>),
where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.
You can obtain their contact details by sending a fax to (352) 29 29-42758.

European Commission

EUR 26797 EN – Joint Research Centre – Institute for Health and Consumer Protection

Title: Alternative methods for regulatory toxicology – a state-of-the-art review

Authors: A. Worth, J. Barroso, S. Bremer, J. Burton, S. Casati, S. Coecke, R. Corvi, B. Desprez, C. Dumont, V. Gouliarmou,
M. Goumenou, R. Gräpel, C. Griesinger, M. Halder, A. Janusch Roi, A. Kienzler, F. Madia, S. Munn, M. Nepelska, A. Paini,
A. Price, P. Prieto, A. Rolaki, M. Schäffer, J. Triebe, M. Whelan, C. Wittwehr, V. Zuang

Luxembourg: Publications Office of the European Union

2014 – 461pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-39651-9 (PDF)

doi:10.2788/11111

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

doi:10.2788/11111

ISBN 978-92-79-39651-9

