

Comments on “CLH report: Proposal for Harmonised Classification and Labelling: 2-(2H-benzotriazol-2-yl)-p-cresol (UV-P)

The members of the European Light Stabilisers and Antioxidants Association (ELISANA), a Sector Group of Cefic, are pleased to provide scientific comments on the CLH proposals for 2-(2H-benzotriazol-2-yl)-p-cresol (UV-P) on the hazard classes skin sensitisation and hazardous to the aquatic environment.

Substance: 2-(2H-benzotriazol-2-yl)-p-cresol

CAS number: 2440-22-4

EC number: 219-470-5

Executive summary

Conclusion on the Evaluation of Environmental Hazards

- As reliable, relevant, and valid experimental data on the degradability of the substance are available, the estimated data (calculated with BIOWIN v4.11) provided by the Dossier Submitter should not be used for the assessment and should be removed from the CLH report.
- Experimental as well as QSAR bioconcentration data demonstrate that the BCF is < 2000 L/kg and thus not bioaccumulative.
- The BCF for UV-P alone is likely to be < 500 L/kg as indicated by comparison of the three bioconcentration studies with different analytical methods: determination of total radioactivity (no distinction between parent, metabolites, and assimilated carbon) vs. substance-specific chemical analysis.
- Metabolism in organism is supported by studies with oral uptake by rats and humans, which show a rapid uptake with a subsequent rapid metabolization in the liver and rapid excretion via the kidney.
- For the assessment of long-term aquatic hazard, the critical value is the 21-d NOEC derived in the Daphnia magna reproduction test according to OECD TG 211 (BASF SE, 2020). In contrast to the value considered by the Dossier Submitter (21-d NOEC = 0.0083 mg/L, measured), the Registrant demonstrated that UV-P remained stable in the exposure system and was not lost due to degradation, volatilization or significant adsorption to the test vessel.
- The Registrant is therefore of the opinion that it is justified to consider the nominal effect value of the study: 21-d NOEC = 0.013 mg/L.
- Based on the 21-d NOEC of 0.013 mg/L, the substance is to be classified as Aquatic Chronic 2. An M-factor of 1 is applicable.

Conclusion on the Evaluation of Human Health Hazards

- The Dossier Submitter (DS) proposes a harmonized classification for Skin Sensitization as Category 1 based on the data available since “2-(2H-benzotriazol-2-yl)-p-cresol acts as a skin sensitiser as shown by human data. There are no OECD TG-conform and reliable animal data available to conclude on the potency of 2-(2H-benzotriazol-2-yl)-p-cresol and therefore, available data do not allow for sub-categorisation.”
- The registrants disagree with this conclusion and will provide evidence on the possibility of sub-categorization for skin sensitization in this document.
- According to Regulation (EC) 1272/2008 (CLP), Annex I, section 3.4.2.2.1.1, “Skin sensitizers shall be classified in Category 1 where data are not sufficient for sub-classification.” Annex I, section 3.4.2.2.1.2 specifies that “Where data are sufficient a refined evaluation [...] allows the allocation of skin sensitisers into sub-category 1A, strong sensitisers, or sub-category 1B for other skin sensitisers.”
- The registrants have classified the substance under evaluation as Skin Sens 1B based on a weight-of-evidence approach, laying down the criteria specified in CLP and ECHA’s Guidance on the Application of the CLP criteria, taking into consideration all available and reliable data. In the following sections, the weight-of-evidence approach will be presented in detail.
- Overall, both animal data and human data available support the criteria in CLP laid down for Skin Sens 1B. Thus, sub-classification as requested under CLP can be performed and should be applied to result in classification of 2-(2H-benzotriazol-2-yl)-p-cresol as Skin Sensitiser, Category 1B.

General comments

Please consider the relevant information as retrieved from latest REACH update Dossier as submitted by 08 March 2024 (tbc) for your evaluation.

ABOUT ELISANA

The European Light Stabilisers and Antioxidants Association (ELISANA), a sector group of Cefic, was created in 2004 with the mission to become the trusted reference on health, safety and environmental information related to antioxidants and UV light stabilisers.

Evaluation of Environmental Hazards

11.1 Rapid degradability of organic substances

11.1.1 Ready biodegradability

2-(2H-Benzotriazol-2-yl)-p-cresol is not readily biodegradable as no biodegradation was observed in two ready biodegradability screening tests. In the study according to OECD 301B, the substance was degraded to 2% based on CO₂-evolution (Ciba-Geigy Ltd., 1989). In a study performed by the Japanese authorities, the degradation was similarly low. The study was performed according to OECD TG 301C. The substance was degraded to 1% based on oxygen consumption and to 2% based on substance specific chemical analysis (NITE, 2023).

The Dossier Submitter (DS) provided additional estimated data on the biodegradability of 2-(2H-benzotriazol-2-yl)-p-cresol using the model BIOWIN v4.11. However, there is no need for the estimation of the biodegradability of the substance as reliable, relevant, and valid experimental data are available. The assessment of the substance as not readily biodegradable is not affected by the estimated data. In addition, the presented data are not fully reliable. The DS already stated in Table 13 (p. 28) of the CLH report that in case of BIOWIN 1 and BIOWIN 2 the training set does not contain benzotriazoles (or similar substances).

The Registrant therefore suggests removing estimated data as they are not relevant for the assessment of the (rapid) biodegradability of the substance. This procedure has also been applied by the Dossier Submitter in case of the estimated bioaccumulation (Ch. 11.3.1, p. 31).

11.3 Bioaccumulation

Table 14 of the CLH Dossier lists the considered information on bioaccumulation by the Dossier Submitter. The Registrant would like to point out that besides the experimentally determined log K_{ow} and the two bioconcentration studies with fish, toxicokinetic data determined in rats and humans are available, which provides details on how UV-P is metabolized and excreted in various vertebrates. The updated IUCLID Dossier contains additional information on the bioaccumulation potential of UV-P: QSAR estimations and experimental data on the bioaccumulation in zebrafish (Zhang et al., 2021). The information is described and discussed in the following chapters.

Toxicokinetic aspects of bioaccumulation

For the assessment of the bioaccumulation potential of UV-P, two studies investigating the bioaccumulation of the substance in fish were performed according to OECD TG 305: 1) BASF SE (2020), 2) NITE (1998). The study by BASF SE (2020) was performed with radiolabelled test material and measuring total radioactivity; thus, the determined BCF values do not distinguish between the parent substance UV-P, any metabolites, and assimilated (radioactive) carbon. In contrast, the study by NITE (1998) was performed with parent-specific chemical analysis, which leads to a BCF representative for the parent substance alone. Metabolites were not investigated. The differences in the analytical procedures lead to different BCF values. Therefore, the BCF based on total radioactivity (BCF_{kgL} = 1456 L/kg) overestimates the bioaccumulation, which can be seen, when comparing the data with the parent specific BCF determined by the Japanese authorities (BCF < 500 L/kg).

For the aspect of metabolizing UV-P in the assessment of the bioaccumulation potential the Registrant would like to present toxicokinetic data for UV-P. Two studies with rats are available which are suitable for an insight into metabolism and elimination of the substance: Ciba-Geigy Ltd. (1977a) and Ciba-Geigy Ltd. (1977b). One study investigated the effects of UV-P on the hepatic xenobiotic metabolism enzymes in the rat. The other study focused on the distribution and elimination of UV-P in the rat. In both studies the substance was administered via oral gavage. Details of these studies are given in the Annex to these comments. The following conclusions can be drawn from these toxicokinetic studies:

- After oral administration of UV-P, the substance is rapidly taken up after oral administration, rapidly metabolized by the liver and excreted via the kidney.
- 91% of a single UV-P dose administered was eliminated from the body within 2 days (94% after 7 days).
- After single administration, 69% of the dose is excreted via the kidney and 25% is excreted through the faeces.

In addition, the toxicokinetic behaviour of UV-P after oral application was also investigated in humans (Fischer et al., 2023). The analysis of blood and urine samples demonstrated a fast resorption of UV-P with a maximal blood level after 1 to 4 h and a maximal elimination via urine after 2.6 to 5 hours. The renal excretion rate is $2896 \pm 885 \mu\text{g/h}$ for UV-P. Based on the available data, it can be concluded that UV-P is not likely to bioaccumulate in both rats and humans.

For the assessment of bioaccumulation in aquatic organisms, metabolism in fish is also expected, which is supported by the large gap between the experimentally determined BCF values determined by BASF SE (2020) and NITE (1998).

Bioaccumulation study: BASF SE (2020)

The Registrant agrees with the fact that the experimentally determined BCF for UV-P in the study by BASF SE (2020) was determined to be $> 500 \text{ L/kg}$. However, in the experimental study by BASF SE (2020) referred to in Table 14 (p. 27) discrete BCF values were determined, which should be reported clearly in this overview. In the text of Chapter 11.3.2 (p. 28), the BCFSS is given as 1622 L/kg and the BCFKgl as 1471 L/kg . These values should be stated in the overview table.

In addition, it should be noted that the determined BCF is based on total radioactivity. As this method does not distinguish between the parent substance, metabolites, and assimilated radioactive carbon, the measured BCF is expected to be higher than it would be based on the substance alone.

The DS however does not accept the experimentally derived BCF values. The DS recognizes that the BCFSSL and the BCFKgl are similar, but the DS disagrees with the study report that steady state was reached. Therefore, the DS judges the BCFs as not reliable. With regard to the kinetic data (uptake rate constant, k_1 and depuration rate constant, k_2), the DS questions the reliability of the k_1 and only accepts the k_2 as reliable. Although k_1 is within the range of model expectations, k_1 is not accepted by the DS due to a fluctuation and a large confidence interval resulting from an unexpected increase in the fish concentration just before the start of the depuration phase.

Instead of the experimentally derived BCF, the DS applied the OECD BCF Estimation Tool (v.2) to derive a kinetic BCF. Details of the estimates are not given except for a statement that 11 out of 14 calculated BCF values were greater than 500. The Registrant cannot accept the results of this

estimation tool especially with regard to the undefined uncertainties of the estimated results. Details for the rejection are described in the following paragraphs.

Comments on the usability of the OECD BCF Estimation Tool (v. 2)

The Dossier Submitter used the OECD BCF Estimation Tool (v.2) for estimating a BCF, which is based on the depuration rate constant (k_2) alone independent of the experimentally determined uptake rate constant (k_1). The tool is intended for the estimation of a BCF from bioconcentration studies with dietary exposure, which results in biomagnification factors (BMF). This “transformation” is required as in most regulatory contexts the BCF is the measure for an assessment of the bioaccumulation potential of a substance and not the BMF. The approach is also referred to in OECD TG 305 (OECD 2012) and in the recently updated ECHA Guidance Document R.11 (R.11.4.1.2.3, ECHA, 2023). It is specifically stated that these approaches are for the estimation of tentative BCFs from data collected in a dietary exposure study. The OECD guideline does not consider the estimation of the uptake rate (k_1) or BCF from studies with aqueous exposure. Further settings and limitations of this tool are indicated below.

The main assumption of the tool and its methods is the independency of the depuration from the uptake route. Therefore, the tool uses only the depuration rate constants k_2 with regard to measured kinetic study data. The following experimental data are needed as well by the tool: growth rate, duration of uptake and depuration phase, and the fish lipid content. Depending on the equation used, either log K_{ow} , fish weight, or both are required for deriving the uptake rate constant k_1 . While fish weight has been easily measured during the study (BASF SE, 2020), it may sometimes be difficult to determine a reliable and valid log K_{ow} . In case of the UV-P, two reliable and very similar experimental values are available from a guideline study (Ciba-Geigy Ltd., 1988) and a reliable database (SRC database: Hansch et al., 1995).

Three methods are included in the tool. Method 1 is the “Uptake rate constant estimation method”, which is described in Annex VIII of OECD TG 305. Thirteen different allometric regression equations and QSAR-like approaches are available for this method. The limitation of the equations is the limited information available on their applicability domains. For the equation of Sijm et al. (1995), a log K_{ow} range is given (log K_{ow} : 3.6 to 8.3), while the information on the other models is rather vague. According to §239 of OECD GD 264 (OECD, 2017), the assumed applicability domain for these models can be estimated from the more detailed information included by Sijm et al. (1995) and the information on substance types included in Barber’s (2003) reanalysis of models. In OECD GD 264, it is concluded that this approach should be useable for aromatic hydrocarbons, those that are chloro-, bromo-, nitro- substituted, and may be suitable for organochlorine and organophosphate pesticides, triarylphosphates and alcohol ethoxylates with log K_{ow} in the range around 3.5 – 8.5. The authors point out that particular care must be taken when using these equations for larger, or higher molecular weight molecules where there is an indication that uptake may be over-predicted. However, no suitable range for the molecular weight is given. Compared to other QSAR models, e.g., the different models of the Catalogic program suite by OASIS, a well-defined applicability domain is not available.

With regard to the publication date of the equations in the OECD BCF Estimation Tool (method 1: 1980 to 2003), it can be assumed that the training sets have been derived from non-standardized bioconcentration tests, which would not fulfil the conditions of the current guidelines. Further, it remains unclear if the models have been validated by testing the equations with separate datasets. It should be noted that the Dossier Submitter has rejected the experimentally derived BCF by NITE (1998) as it does not fulfil the requirements of the current OECD TG 305 (OECD 2012).

The OECD GD 264 also mentions that an estimation of the uncertainty is not possible for the BCF estimates. Estimates of uncertainty in the predicted parameters (uptake rate constant and BCF) are not possible, because they are related both to the (dietary) study measured data and the models used in the prediction, including their underlying training sets. If the estimated BCF values are close to or just above a critical BCF level, i.e., 2000 L/kg or 5000 L/kg, a scientific evaluation of the bioaccumulation potential and its reliability based on the uncertainties of the predicted BCF cannot be performed.

Annex VIII of OECD TG 305 refers to the review of Brooke and Crookes (2012). The authors state that “no one method is more correct than the others.” They point out that a clear justification should be given for the model used. The choice of model may be influenced by the level of validation and applicability domain. However, as information on validation and applicability domain are missing, this step is not practicable. Nevertheless, for the prediction of BCF values, the used/selected models should be justified. An evaluation should be performed as a weight-of-evidence approach including the k1 estimate, the estimated BCF, the experimentally derived BMF, and other substance parameters (e.g., log Kow). In case of the applied calculations performed by the Dossier Submitter, the experimentally determined BCF should have been considered as well, especially as the studies fulfil the validity criteria of OECD TG 305.

The Dossier Submitter applied the tool without a justification for the usability and reliability of the models as required by REACH Annex XI section 1.3 and the generally scientifically accepted OECD QSAR assessment framework. There was no comment on the applicability domain for the individual models. Instead, the Dossier Submitter used all available methods without a further evaluation or comparison of the results by simply stating that “the majority” of the BCF values was “>> 500”. The Registrant does not agree with this approach, especially as a valid experimental result is available.

Bioaccumulation study: NITE (1998)

Table 14 (p. 27) states lipid-normalized BCF values for two test concentrations 0.1 and 0.01 mg/L. The Dossier Submitter performed the lipid-correction based on a lipid content of 3.6%. However, the original study reports separate lipid contents for the applied concentration levels. At 1.0 and 0.1 mg/L, the fish had a lipid content of 3.6%, while the lipid content at the lowest concentration level was 4.0%. Therefore, the lipid-corrected BCF values at 0.01 mg/L are lower than stated in Table 14 of the CLH Dossier: 55–275 L/kg.

Concentration level	Lipid content	Reported BCF (NITE (1998): min–max	Lipid-normalized BCF (5% lipid):
0.1 mg/L	3.6%	130–295 L/kg	181–410 L/kg
0.01 mg/L	4.0%	44–220 L/kg	55–275 L/kg

The Registrant recognizes that the study does not fulfil the requirements of the currently applicable OECD TG 305 (OECD 2012); however, the results are sufficiently reliable to be used as supporting data. The Dossier Submitter refers to the measurement of the lipid content only at the beginning of the uptake phase and not at its end as well. However, the lipid contents at test start were at 3.6% and 4.0% depending on the concentration level, which is close to the 5% level for normalization. The original BCF values as well as the lipid-normalized BCF values are stated in the table above. The corrections did not lead to drastic changes of the BCF values. The lipid content of the fish at the end

of the uptake period is most likely close to the start values as the fish were regularly fed with an appropriate diet during the study. The lipid content would have to have dropped to less than 1% for the BCF to exceed the critical value of 2000 L/kg for B substances. In this case, it can be assumed that the study would not be valid as the fish would not have been fed properly and would be starving.

The Dossier Submitter stated that it is unclear if the water quality was maintained within acceptable limits over the entire test due to missing details in the summary report. The Japanese study report contains the missing information, e.g., range of the dissolved oxygen concentration at the three investigated concentration levels. In addition, the study was performed with flow-through conditions using quality-checked ground water as test medium, daily cleaning of the test vessels, and regular measurement of the dissolved oxygen concentration. It can be therefore assumed that the water quality was acceptable, which should not question the reliability of the study.

As already pointed out above, the bioaccumulation was investigated by applying parent-specific analytical methods. Radio-labelled test material was not used. Therefore, in contrast to the study by BASF SE (2020), the BCF determined by NITE (1998) does not consider metabolites nor assimilated carbon from the substance. Thus, the fact that the BCF in the Japanese study is lower than the BCF in the BASF study is due to the different analytical methods and not due to deficiencies of the studies themselves.

Bioaccumulation study: Zhang et al. (2021)

A reliable publication of Zhang et al. (2021) supports the findings on the toxicokinetic as well as the bioaccumulation potential as determined by BASF SE (2020) and NITE (1998). Zhang et al. (2021) exposed the adult female *Danio rerio* (43±1 mm, weight 442±34 mg, age 3-4 months) to 0.5 and 10 µg/L of the test item according OECD guideline 305 for 28 days to investigate the tissue specific accumulation potential. Fish were fed daily with commercial feeds at a rate of 1.0% body weight. After the uptake phase a 14-d depuration phase without the test item followed. The BCF values, the depuration half-lives and kinetic parameters were determined organ specific (liver, muscle, kidney, ovary, intestine, skin, gill). BCF values could not be derived for the lower test concentration as the measured tissue concentrations were below the limit of quantitation (LOQ = 9.26–31.1 ng/g ww). BCF values between 10.4 and 667 L/kg were determined for the different tissue types. The highest value was determined for the intestines, the lowest in the ovary. Although the BCF values were not growth corrected, based on the size of the fish and the feeding rate, it can be concluded that no significant growth occurs during the experiment. The results were also not lipid normalized, however, due to the use of adult female fish only, an overall higher lipid content, compared to juvenile or male fish can be assumed. Measured lipid content was highest in ovaries (22% ww), kidneys (17% ww) and liver (8% ww). The organ weights were 0.519, 1.51, 2.94, 3.4 and 5.74, and 12.5% for kidney, gills, intestines, liver, skin, and ovary, respectively. Thus a lipid normalization would even lead to a reduction in measured BCF values in these organs and may even lead to a reduced lipid corrected BCF for the whole fish. The determined depuration half-lives between 0.215 (intestine) and 1.87 d (skin) are short and support the results of the toxicokinetic investigations. Liver tissue was selected by the authors for detecting and identifying biotransformation products. Two metabolites were identified for UV-P.

Similar to the study by NITE (1998), the test item was not radio-labelled, and substance-specific chemical analysis was performed by Zhang et al. (2021); therefore, the BCF values refer to the parent only. The relatively low BCF values of Zhang et al. (BCF ≤ 667 L/kg) are in agreement with the values of NITE (BCF ≤ 410 L/kg). Since the BCF from the study by BASF SE (2020) refers to a combined

value for parent and products based on the measurement of radioactivity, the determined lower BCF values for the parent within this publication are assumed to be plausible. In conclusion, low BCF values and fast depuration times support the conclusion that the substance does not fulfil the B/vB criteria.

Supporting QSAR calculations

The experimental results are supported by reliable QSAR calculations. The BCF base-line model considers mitigating factors like metabolism, molecular size and the water solubility when calculating the BCF. The maximum BCF was estimated to be 1361 L/kg, which is in line with the BCF of the study by BASF SE (2020). In both cases, no differentiation between parent and metabolites is made, leading to a higher BCF. In contrast the BCF is significantly reduced when mitigating factors are considered in the calculation process. Thus, the final BCF is estimated to be 42 L/kg, indicating a low potential for bioaccumulation.

The substance is within the parametric and the metabolic domain of the model. However, it is outside of the structural domain. 83% of the atom-centred fragments were found in correctly predicted training set chemicals. The other 17 % of the fragments were incorrect fragments. Although this lowers the confidence the model prediction is regarded as reliable indication.

The predicted BCF values using the BCFBAF model v3.01 of EPI Suite v4.11 support the conclusion on a relatively low bioaccumulation potential: BCF = 324 and 364 L/kg. The regression model considers the log Kow of 4.20 in the calculation of the BCF. A similar value is calculated by the Arnot-Gobas model for the mid-trophic level with biotransformation. For both models, the substance is in the applicability domain of the models.

Model	BCF (L/kg)	Remarks
OASIS Catalogic v5.15.2.14: BCF base-line model v05.12	42	All mitigating factors considered (BCF _{max} = 1361 without mitigating factors); Completely in parametric and mechanistic applicability domain; all relevant structures within AD (83% of total structures))
EPI Suite v4.11: BCFBAF model v3.01; regression-based estimate	324	Within the applicability domain
EPI Suite v4.11: BCFBAF model v3.01; Arnot-Gobas BCF & BAF methods	364	Mid trophic level, including biotransformation; Within the applicability domain

11.5.2 Chronic toxicity to aquatic invertebrates

The study was performed according to OECD TG 211 with semi-static exposure. The test solutions were prepared from a concentrated solution in the solvent dimethylformamide (DMF) with a final solvent concentration of 100 µL/L. Renewal of the test solutions was performed three times per week (every 48 and 72 h, respectively). Food algae were administered daily according to OECD TG 211 (Table 2). The test concentrations were analytically monitored in fresh and aged test solutions with and without test organisms (and food algae). The monitored test solutions were from the following renewal periods:

- Day 2–4: with organisms
- Day 11–14: with and without organisms
- Day 16–18: with and without organisms

The mean measured values of the initial test concentrations were within an acceptable range of 100 to 113% of nominal concentrations (Table 1). The aged test solutions (48 to 72 h) with *Daphnia* showed significantly lower concentrations: 31 to 45% of the corresponding initially measured values (Appendix I). An exception is the highest test concentration (0.130 mg/L), which was at 82% of the mean initial values (Table 1). Based on the time-weighted mean, the test concentrations ranged between 53 to 64% of the nominal values or 47 to 57% of initially measured concentrations (0.0013–0.041 mg/L). In case of the highest test concentration the recovery was 93% and 83%, respectively. In the replicates without organisms, no significant loss of the test substance was observed. The time-weighted mean concentrations ranged from 101 to 117% of the nominal values or 96 to 101% of initially measured concentrations (Table 1).

Table 1: Measured test substance concentrations in samples with daphnids/algae (renewal periods: days 2–4, 11–14, and 16–18; 48–72 h old) and without daphnids/algae (renewal periods: days 11–14, and 16–18; 48–72 h old)

Nom. conc.	2-(2H-benzotriazol-2-yl)-p-cresol							
	Initial		With daphnids/algae			Without daphnids/algae		
			Time-weighted mean			Time-weighted mean		
[µg/L]	[µg/L]	[% of nom.]	[µg/L]	[% of nom.]	[% of meas. mean initial]	[µg/L]	[% of nom.]	[% of meas. mean initial]
Control	0	-	0	-	-	0	-	-
Solvent control	0	-	0	-	-	0	-	-
1.3	1.3	100	0.72	56	56	1.31	101	101
4.1	4.6	109	2.5	61	56	4.35	106	99
13.0	14.2	112	8.3	64	57	14.4	111	98
41	38.6	113	21.6	53	47	47.9	117	96
130	133	112	120.4	93	83	145.7	112	96

Since the substance is neither considered as biodegradable nor as volatile, the “loss” of the test substance in the exposure vessels was likely due to adsorption of the test material to food algae and daphnids. According to the requirements of OECD TG 211, the juvenile daphnids receive between 0.10 to 0.14 mg TOC/animal/day, while the adult daphnids receive about twice as much food from day 8 to the end of the exposure phase (0.20 mg TOC/animal/day; Table 2). The influence on the algal biomass on the analytical recovery in the water phase can be derived from the decreasing recovery rates with time. The aged water samples in the initial phase of the study (day 2–4) show a higher recovery than the aged water samples of the samplings later in the exposure phase (day 11–14 and 16–18; Appendix I). It can be summarized that with an increase in the applied amount of food algae, the recovery of the test substance in the test solutions decreased.

The detected loss of the test material from the aqueous phase of the test system is addressed in OECD TG 201 and OECD GD 23. In OECD TG 201 (Algal growth inhibition test, section 40), it is specified that “disappearance of the test substance from solution by adsorption to the increasing algal biomass does not mean that it is lost from the test system.” The aspect of adsorption of the test substance to food algae in chronic daphnid tests is considered in OECD Guidance Document 23 (OECD, 2019): “Adsorption may also be a problem in chronic daphnid studies where test chemical adsorbed to the food algae can lead to apparent reduction in the freely dissolved concentration (when algae are separated prior to analysis) but would still provide a secondary exposure route via ingestion.” It should be mentioned that the amount of algae given as feed to the daphnids according to OECD TG 211 is not directly comparable to the algae concentrations used in OECD TG 201. The amount of food algae in OECD TG 211 is specified as total organic carbon (TOC/animal/day) and not as cell number as in OECD TG 201 (cells/mL). However, the mechanism of adsorption is identical.

As described in both OECD documents, adsorption to algae is not identical with a loss from the test system. In order to document the amount of algae in the test system within one renewal period of 72 hours, photos were taken every 24 hours before and after applying the food algae to the beakers (Appendix II). The observed animals were from the GLP in-house daphnid culture representing the same culture conditions (renewal period, amount of algae fed; Table 2), as used in the before-mentioned OECD TG 211 GLP guideline study. The following three scenarios were monitored: without daphnids, with juvenile daphnids, and with adult daphnids (Table 3). The amount of algae fed daily was in accordance with the requirements of OECD TG 211 (Table 2). The photos were taken from above looking onto the water surface; therefore, it is possible to get an impression on the algae concentration in the vessels by the differing intensity of the green colour of the culture medium. Figure 1 and 2 show selected pictures after 48 and 72 h from a time series as available in Appendix II. In all scenarios, an increase in the green colour of the test medium with time could be observed. The beakers without daphnids show a continuous increase in green with each addition of the algae, which can also be observed in the vessels with daphnids but with a far paler tone.

Table 2: Feeding schedule according to OECD TG 211 and amount of food (TOC) given to cultured daphnids

Day of Test	Amount of food [mg TOC/animal/day]*	Amount of food TOC [µg algae concentrate/animal/day]
0 – 3	0.10	38
4 – 5	0.12	46
6 – 7	0.14	54
8 – End	0.20	77

*Amount of feed algae determined as total organic carbon

Table 3: Overview of culture beakers with and without daphnids

Beaker no.	Description of the scenarios
1	Juvenile daphnid (0-3 days of age), fed daily with algae
2	Test medium without daphnid, food algae applied identical to beaker #1
3	Adult daphnid (25-28 days of age), fed daily with algae
4	Test medium without daphnid, food algae applied identical to beaker #3

The comparison of beakers with daphnids after 48 h and 72 h before feeding with those without daphnids demonstrates the effect of the daphnids feeding on the algae. It can be easily derived that the amount of algae administered to the beakers is quantitatively taken up by the daphnids within the 24 h feeding interval (Fig. 1, Fig. 2, Appendix II). The algal biomass remaining after a 24 h interval is very low, which is well observable in the beakers with the adult daphnid (Appendix II).

Overall, the pictures support the assumption that the majority of the food algae are taken up by the daphnids during an OECD TG 211 study using a semi-static test design with 48 and 72 h renewal periods, respectively. Although a small fraction of the algae may remain in the test vessel uneaten, any substance adsorbed to the food algae is likely being eaten by the daphnids and thus cannot be considered as lost from the system but rather provided an additional exposure route to the test organism.

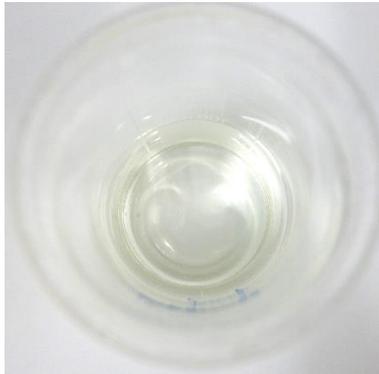
The adsorption of the test substance to algae is also demonstrated in the growth inhibition test with algae (BASF SE, 2018). The measured initial test concentrations were in the same range as in the Daphnia reproduction study (BASF SE, 2020) ranging from 0.0006 to 0.135 mg/L (initially measured). The test solutions were prepared without solvent. Instead, a saturated solution was prepared by the slow-stirring method, which was checked for the absence of undissolved test material. Test vessels were pre-conditioned by exposing them with the test solutions for 72 hours, which should have reduced the potential for adsorption of the test substance to the test vessels as suggested by Section 106 of OECD Guidance Document 23 (OECD, 2019). The analytical monitoring showed that the test concentrations in the aqueous phase were significantly lower than the initial values (< LOQ to 37% of initially measured concentrations). However, test solutions which were exposed under the same conditions but without algae, showed concentrations in the four highest test solutions. For the two lowest test concentrations (0.1 and 0.316 mg/L), the recovery without algae was still clearly higher than in the test assays with algae (with algae: ≤ 10%; without algae: 57 and 63%).

In summary, the test substance is partially adsorbed by the food algae provided to the daphnids. Therefore, the daphnids were exposed to the test substance dissolved in water, as well as to the test substance adsorbed to the food algae via ingestion and to the test substance adsorbed to the daphnids themselves. It can be concluded that no actual loss of the test substance from the exposure system occurred. As the test substance was stable during the test and the recovery remained in the range of ±20% of nominal concentrations as demonstrated in replicates without daphnids and algae, it is justified to derive the effect concentrations from the nominal test substance concentrations.

The 21-d NOEC for reproduction as well as for mortality is 0.013 mg/L based on the nominal concentrations (Table 4).

Beakers after 48 h containing M4-medium with 0.1 mg TOC/algae/day added

With young daphnid

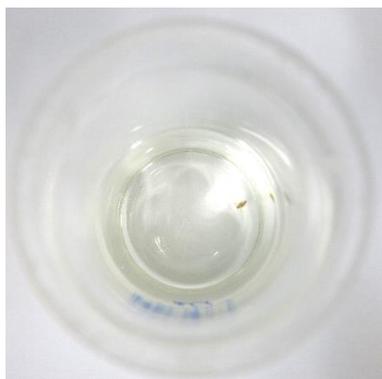


Without young daphnid



Beakers after 48 h containing M4-medium with 0.2 mg TOC/algae/day added

With adult daphnid



Without adult daphnid



Figure 1: Beakers with young and adult Daphnids and feed algae after 48 hours

Beakers after 72 h containing M4-medium with 0.1 mg TOC/algae/day added

With young daphnid



Without young daphnid



Beakers after 72 h containing M4-medium with 0.2 mg TOC/algae/day added

With adult daphnid



Without adult daphnid



Figure 2: Beakers with young and adult Daphnids and feed algae after 72 hours

Table 4: Reproduction and mortality of daphnids after 21 days

Test groups	Reproduction		Mortality			
	Mean living young per surviving adult	% Effect ^b	Parent animals	% Effect ^b		
0 (Control)	112 (14.3% ^a)	-	10	-		
0 (Solvent)	113 (9.4% ^a)	-	10	-		
0.001 / 0.0007	111	-	10	-		
0.004 / 0.0025	115	-	10	-		
0.013 / 0.0083			109	-	9	10%
0.041 / 0.0216			0	-	0	100%**
0.130 / 0.120			0	-	0	100%**

**Statistically different than the control group at $p \leq 0.01$

a: Coefficient of variation for control fecundity

b: Effect relative to control. Only calculated for statistically significant effects.

11.5.3 Chronic toxicity to algae or other aquatic plants

As described above, the measured test concentrations at the end of the exposure (72 hours) were reduced by adsorption to the test organisms (Table 5). This is demonstrated by the comparison of the measured test substance concentrations in replicates with and without algae. The additional analysis of abiotic replicates of the test levels without algae is described in OECD Guidance Document 23 (OECD, 2019) as a method to better characterize exposure levels. Since the substance is neither considered as biodegradable nor volatile, the “loss” of the test substance in the exposure vessels was likely due to adsorption of the test material to the algae. Adsorption to the test vessel can be considered insignificant as the test vessels were pre-conditioned with the test solutions prior to the exposure of the algae as suggested by Section 106 of OECD Guidance Document 23 (OECD, 2019). Thus, there was no actual loss of the test material from the test system.

This approach is in agreement with the conditions described in OECD TG 201. The guideline describes in section 40 that the alga growth inhibition test is a more dynamic test system than most other short-term aquatic toxicity tests. For adsorbing substances, such as 2-(2H-benzotriazol-2-yl)-p-cresol, the actual exposure concentrations may be difficult to define at low test concentrations. Therefore, disappearance of the test substance from solution by adsorption to the increasing algal biomass does not mean that it is lost from the test system.

Consequently, the effect values could also be expressed based on measured initial concentrations of the abiotic samples. However, the relevant effect concentrations (72-h ErC10 = 0.059 mg/L; 72-h ErC50 > 0.082 mg/L) for the classification have not been recalculated based on measured initial concentrations of the abiotic samples as they are already higher than the critical levels of the CLP regulation and would not influence the classification.

Table 5: Measured test substance concentrations in samples with and without algae

	2-(2H-benzotriazol-2-yl)-p-cresol				
Sampling date	0 h	72 h		72 h	
Nominal loading of the test item	Initial	With algae		Without algae	
[mg/L]	[µg/L]	[µg/L]	[% of meas. initial]	[µg/L]	[% of meas. initial]
10.0	135	50.2	37	135	100
3.16	46.6	6.54	14	43.5	93
1.00	14.1	2.99	21	13.6	96
0.316	4.64	0.534	12	3.75	81
0.100	2.08	0.21	10	1.18	57
0.0316	0.558	< LOQ		0.35	63
Control	< LOQ	< LOQ		< LOQ	

Evaluation of Human Health Hazards

Available animal data

Several animal studies are available with the test substance. However, based on the experimental nature of many of those studies, only two studies are considered relevant in the context of determination of potency. Therefore, only those two studies are further described in detail in this section. Nevertheless, all available animal studies and their respective outcomes are listed in Table 8.

Guinea pig maximization test (Ciba-Geigy Ltd., 1992)

The test substance (2-(2H-benzotriazol-2-yl)-p-cresol) was assessed for its skin sensitizing properties in a guinea pig maximization test (GPMT) according to OECD guideline 406 and under GLP (Ciba-Geigy Ltd., 1992). Male and female Pirbright White guinea pigs (10 per sex and group) were used, and induction was performed by first intradermal injections with the test substance (> 98% purity). Three pairs of intradermal injections with 0.1 mL each were performed simultaneously into the shaved neck of the guinea pigs (adjuvant/saline 1:1; test substance in peanut oil; and test substance in adjuvant mixture). A control group of five animals per sex was treated with adjuvant and vehicle alone during the induction period. In the second week, the test substance was incorporated into Vaseline and applied on a filter paper patch to the neck of the animals as occluded administration for 48 hours. Two weeks later, challenge was performed with test substance or vehicle in Vaseline, which was applied to the flank under occluded conditions for 24 hours. Test substance concentrations were evaluated prior to study conduct on separate animals. Intradermal induction was performed with 5% test substance in peanut oil as it could be injected and was well tolerated. Epidermal applications were conducted with 30% test substance in Vaseline based on a preliminary experiment, which showed that 30% test substance application did not cause irritant reactions. For the challenge, 20% in Vaseline was used as sub-irritant concentration, since higher concentrations may lead to nonspecific reactions in adjuvant-treated animals (Magnusson, 1980).

Animals induced with vehicle and challenged with vehicle alone did not show any reaction. Similarly, those animals induced with test substance and challenged with vehicle alone did not show any reactions either 24 or 48 hours after removal of the dressings. Animals induced with vehicle but challenged with 20% test substance in Vaseline did show some reactions after 24 and 48 hours (10% and 20% with reactions, respectively). The skin changes were however rather weak; erythema score was maximally 1 and no oedema was observed. The group induced with test substance and challenged with test substance showed reactions in 80% of animals after 24 hours and 90% of animals after 48 hours (Table 6). After 24 hours, the reactions scored in animals with skin changes included erythema grades 1-3 and oedema 0-2. After 48 hours, erythema grades 1-2 and oedema grades 0-2 were noted for affected animals.

The sensitivity of the test system was assessed in periodical positive control tests, which showed 100% sensitization for positive control substance 1-Chlor-2,4-dinitrobenzol and 0% sensitization rate in concurrent control animals.

Table 6: Results of OECD TG 406 study (GPMT) with 2-(2H-benzotriazol-2-yl)-p-cresol (Ciba-Geigy Ltd., 1992)

	Induction	Challenge	Readout [# positive animals / %]
--	-----------	-----------	----------------------------------

			24 hours		48 hours	
Control group	Vehicle	Vehicle	0/10,	0%	0/10	0%
	Vehicle	Test substance	1/10	10%	2/10	20%
Test group	Test substance	Vehicle	0/20	0%	0/20	0%
	Test substance	Test substance	16/20	80%	18/20	90%

The DS claims that uncertainty on the reliability of the data comes from the 10% and 20% response rates in the control animals challenged with test substance. However, it is well-known that guinea pigs previously treated with adjuvant have lower irritation thresholds in challenge reactions (Frankild et al., 1996; Kligman and Basketter, 1995). In case of doubt on the reliability of results, a re-challenge would be the method of choice to clarify the sensitization potential of a substance. In this case, the substance-treated group was considered clearly positive, therefore a re-challenge was not considered relevant. For the reactions in the control group, a re-challenge might have clarified the nature of the observations. However, based on the very slight reactions observed at low frequency, and based on the literature available on non-specific irritant responses, it can be assumed that the reactions observed in the control group are unspecific irritation rather than specific sensitization. For evaluation of the assay, these reactions are considered of little consequence, as they are only relevant in the context of identification of false-positive reactions. Given the clear sensitization potential observed, the clarification of false positives in the control group was considered not relevant and has no impact on the validity of the data for hazard assessment.

As this study is a guideline-compliant study which was conducted under GLP and the reactions of control animals observed is well-known for this type of study, the results are considered reliable without restriction (Klimisch 1) by the registrants.

For the purposes of classification, CLP Annex I, table 3.4.4 states that for the guinea pig maximization test a substance is considered a skin sensitizer Cat. 1B if $\geq 30\%$ respond at $> 1\%$ intradermal induction dose. Based on the data available, the test substance does fulfil these criteria as Skin Sensitizer Category 1B based on the guinea pig maximization test available.

Local Lymph Node Assay (Ikarashi et al., 1994a)

Two separate tests were conducted in the context of this publication. In a protocol similar to a local lymph node assay (LLNA), 7–10-week-old female Balb/c mice (3 per group) were treated with topical application of 25 μ l of test substance at concentrations of 0.25, 0.5, 1, or 2% in acetone olive oil (AOO) or vehicle alone on the backs of the ears for three consecutive days. On day 5 (4 days following the first exposure), the mice were sacrificed, the auricular draining lymph nodes were extracted, weighed, and processed to single cell suspensions. The cell suspensions were seeded and incubated with ^3H -thymidine for 24 hours *in vitro*. The incorporation of ^3H -thymidine was measured, and a stimulation index was calculated.

None of the concentrations tested led to positive results in this LLNA in two independent experiments at concentrations up to 2% (Table 7).

Table 7: Results obtained in LLNA published by (Ikarashi et al., 1994a)

Experiment	Concentration [%]	Lymph node weight [mg]	³ H thymidine incorporation [mean cpm ± SD x 10 ⁻³]	Stimulation index
1	0	17.3	2.16 ± 0.17	-
	0.25	16.5	2.20 ± 0.19	1.02
	0.5	17.1	3.07 ± 0.18	1.42
	1	14.5	2.19 ± 0.10	1.01
	2	19.1	2.64 ± 0.22	1.22
2	0	n.d.	1.36 ± 0.12	-
	0.5	n.d.	1.07 ± 0.19	0.78
	1	n.d.	1.99 ± 0.24	1.46
	2	n.d.	1.96 ± 0.50	1.44

The test method shows several deviations as compared to the OECD TG 429 protocol. While the animals were of the appropriate species (mice) and similar age (7-10 weeks as compared to 8-12 weeks in OECD TG 429), the mouse strain used differed from the prescription in the OECD guideline (Balb/c vs. CBA strains), however it was in line with the originally described protocol by Kimber and Weisenberger (Kimber and Weisenberger, 1989). The induction protocol was comparable to the OECD guideline, however the analysis of sensitization induction differed in that ³H-thymidine incorporation was conducted *in vitro* instead of *in vivo*. Nonetheless, Kimber and Weisenberger showed with a similar protocol that 4 days after initiation of induction, the proliferation in the draining lymph nodes was near maximum. Therefore, while in deviation from the OECD TG 429, a positive response could have been expected from the protocol as described. In addition, the same protocol was used on known sensitizers such as 2,4-Dinitrochlorobenzene (DNCB), which showed clear increases of stimulation indices, thus showing that the test system is able to detect skin sensitizers (Ikarashi et al., 1993).

Animal numbers used were too low (only 3 per group instead of 4 per group as recommended in the OECD TG 429), however as the experiment was repeated with near-identical results, it is acceptable to combine both experiments. Evaluation of 6 animals per group provides a higher certainty of evaluation and thus also a high validity of the data.

No information was provided on test substance purity.

The dose selection was not in line with that prescribed by the OECD TG 429 as only concentrations up to 2% were tested. Therefore, this publication is not helpful in answering the question if the substance is a sensitizer at all; however, it is helpful to answer whether skin sensitization was observed at 2% in a protocol similar to OECD TG 429.

Based on the deviations from the OECD TG 429 protocol, the low animal numbers used and the dose selection of only up to 2% test substance, this assay was considered reliable with certain restrictions (Klimisch 2). It provides evidence that no skin sensitization was observed at concentrations of 2% in a study similar to OECD TG 429. According to the CLP criteria, LLNA data with EC3 values above 2% allow for classification of substances as skin sensitizing Category 1B. Thus, while this study cannot be

used for classification by itself, it clearly supports the finding that the substance is not a strong skin sensitizer.

Further animal studies with non-standard protocols (Ikarashi et al., 1994b, 1994a, 1993; Yamano et al., 2001)

Several publications are available for modified LLNAs, modified mouse ear swelling tests and modified guinea-pig maximization test.

Numerous shortcomings must be noted for these publications:

- (a) The methods are experimental and not validated
- (b) The test substance purities are not described
- (c) The animal numbers evaluated are low
- (d) Some of the publications are lacking concurrent controls
- (e) The results are often poorly described
- (f) The reproducibility of the assays is unclear
- (g) In some cases, the originality of the data must be questioned as several publications show the same numbers for the same assays.

Based on these experimental weaknesses and the lack of method validation, the interpretation of these data is very challenging and associated with a high degree of uncertainty. Therefore, these data were considered as not reliable (Klimisch 3) for the purposes of classification and considered not suitable for assessment of potency by the registrants.

One publication is available from Lee et al. describing studies which were conducted at least similar to OECD Guideline 406 (Lee et al., 2019). Since these data are only accessible as secondary information, Klimisch 4 rating was assigned.

Summary of animal data available

Table 8: Overview of available (relevant) animal data

Reference	Type of study	Protocol	Reliability (Klimisch)	Sensitizer / non-sensitizer	Supports sub-classification
(Ciba-Geigy Ltd., 1992)	GPMT	OECD TG 406	1	Sensitizer	Cat. 1B
(Ikarashi et al., 1994a)	Modified LLNA	(Kimber and Weisenberger, 1989)	2	Non-sensitizer (up to 2%)	(Cat. 1B)
(Ikarashi et al., 1994a)	Modified MEST	Similar to (Gad et al., 1986)	3	Sensitizer	---
(Ikarashi et al., 1994b)	LLNA with intradermal induction	---	3	Sensitizer	---
(Ikarashi et al., 1994b)	Modified MEST	---	3	Sensitizer	---
(Ikarashi et al., 1993)	LLNA	---	2	Non-sensitizer (2%) – same data as (Ikarashi et al., 1994a)	---
(Ikarashi et al., 1993)	LLNA with intradermal induction	---	3	Sensitizer	---
(Yamano et al., 2001)	GPMT	Deviates from OECD TG 406	3	Sensitizer	---
(Lee et al., 2019)	GPMT	Similar to OECD TG 406	4	Non-sensitizer	---
(Yamano et al., 1993)	Modified GPMT	---	4	Sensitizer	---

Available human sensitization data

Several historical data from testing in healthy humans are available as human repeated insult patch test (HRIPT) or human maximization test (HMT). These data from human induction studies are described in further detail below. An overview of the available data is compiled in Table 9.

In addition, patch tests in the general population as well as patients are available, and the most relevant data are described in detail below. The available information is compiled in Table 10 and Table 11, respectively.

Finally, several publications of case studies are available on the substance, summarized in Table 12.

Human induction studies

HRIPT (Hill Top Research Institute, 1960)

A repeated insult patch test was performed for the substance in 59 volunteers. The subjects received 24-hour patch exposures under occlusive conditions three times weekly over three weeks. In the sixth week, a 24-hour challenge exposure was performed.

The volunteers completing the test were 12 men and 47 women; 9 men and 35 women between the ages of 20 and 50, 3 men and 12 women older than 50 years.

The patch test was a 3/4 x 7/8-inch swatch of "Webril" moistened with 0.5 ml test substance solution (0.5% test substance in dimethyl phthalate). The sample was applied to the upper arm of the subject and left in place for 24 hours. The patch was applied to the same site each time except for the challenge patch, which was applied directly adjacent to the induction site.

None of the volunteers experienced sensitization in this test.

The DS calculated the dose per skin area (DSA) as 230 µg/cm², however this is believed to be due to an error in the submitted RSS, which gave the concentration of test substance as 0.2% in vehicle. Indeed, the materials and methods section of the report as well as the raw data specify the test material was 0.5% in vehicle.

Therefore, the DSA was calculated as follows:

Patch size: 0.75 x 0.825 inches = approx. 1.9 x 2.2 cm = 4.23 cm²

Test substance amount: 0.5 mL of 0.5% solution = 0.0025 g = 2500 µg/application

DSA = 2500 µg ÷ 4.23 cm² = 590 µg/cm²

Based on these data, no sensitization was observed in humans at 590 µg/cm².

This study is considered sufficiently well reported and the study conduct meets generally accepted scientific protocols. Therefore, these data were considered reliable with restrictions (Klimisch score 2) for assessment.

HRIPT (Kligman, 1964a)

A repeated insult patch test was performed on 25 healthy adult males (age range 23-34 years, 16 African American, 9 Caucasian). The induction phase consisted of 5 exposures for 48 hours each with one-day intervals between them. Non-irritating substances (such as the substance 2-(2H-benzotriazol-2-yl)-p-cresol) were used at 25% in petrolatum for induction. Each patch consists of a 1.5-inch square on Webril cloth to which 0.75 g test substance is applied. It was held in place by occlusive dressing.

Each exposure was to the same site on an extremity, which was previously irritated by 24-hour treatment with 5% sodium lauryl sulphate. Two weeks after the last exposure, a new skin area was challenged by a 48-hour patch test with 10% substance in petrolatum. The challenge was applied to a normal area pre-treated for one hour with sodium lauryl sulphate. Evaluation of the challenge site was performed for three consecutive days.

None of the subjects showed an indication of skin sensitization.

Calculation of DSA:

Patch size 1.5 inches (worst-case assumption: 1.5 x 1.5 inches) = 3.81 x 3.81 cm = 14.52 cm²

Test substance amount: 0.75 g of 25% concentration = 187500 µg/application

DSA = 187500 µg ÷ 14.52 cm² = 12916 µg/cm²

The report has several shortcomings, such as a generic description of methods, lack of information on test substance purity, and testing of only 25 subjects. Therefore, this report was assigned a Klimisch 4 rating, but can (in conjunction with other available) data be used to support the conclusion on the evaluated substance.

HRIPT (Kligman, 1964b)

A repeated insult patch test was performed on 25 healthy male adults. The subjects were treated with five 48-hour exposures of 25% test substance in Vaseline (occlusive patches) with one day intervals between exposures. Each treatment site was abraded with sandpaper prior to application of the test substance. Two weeks after the last exposure, a new skin area was abraded and challenged with 10% test substance in Vaseline with a 48-hour occlusive patch. Skin readings were made at 48, 72, and 96 hours.

None of the subjects were sensitized under the conditions tested. Based on the report details available, no DSA could be calculated.

The study has several shortcomings and was assigned a Klimisch 4 rating mainly due to the brief reporting of methods and results, however the data can be used in a weight-of-evidence together with other available information to support the conclusion on the evaluated substance.

Human maximization test (Lee et al., 2019)

A nail polish containing 0.5% drometrizole was applied to the upper back of 148 subjects by topical occlusive patches 3 times per week for 3 weeks. After a two-week break, new patches were applied to untreated sites for 48 hr. Reaction scores were determined at 48 and 96 hr. None of the subjects were sensitized.

As these data were available only as secondary literature, a Klimisch 4 rating was assigned.

Summary human induction studies

Table 9: Overview of human repeated insult patch tests available with 2-(2H-benzotriazol-2-yl)-p-cresol

Reference	Type of study	Reliability (Klimisch)	Applied concentration	# of subjects evaluated	% sensitized	DSA ($\mu\text{g}/\text{cm}^2$)
(Hill Top Research Institute, 1960)	HRIPT	2	0.5%	59	0	590
(Kligman, 1964a)	HRIPT	4	25%	25	0	12916
(Kligman, 1964b)	HRIPT	4	25%	25	0	---
(Lee et al., 2019) and (CIR Expert Panel, 2008)	HMT	4	0.5% in nail polish	148	0	---

Epidemiological study

Cross-sectional Patch test evaluation (Zhao and Li, 2014)

Patch tests were conducted to assess the sensitization prevalence of different substances in the general population. For this purpose, 201 students were recruited via a university. All subjects were of Chinese Han ethnicity; the population included 61 men and 140 women with a mean age of 23.5 ± 2 years (range 19-30 years). Different environmental allergens including 2-(2H-benzotriazol-2-yl)-p-cresol (1% in petrolatum) were assessed with patch tests using the IQ Chamber (Chemotechnique Diagnostics – dimensions: 64 mm^2 , $32 \mu\text{l}$ volume). None of the participants showed a positive response to the substance under evaluation.

As the reporting was considered adequate and the method is widely accepted and standardized in use, the data were considered reliable with restrictions (Klimisch 2).

Summary epidemiological studies

Table 10: Overview of epidemiological patch tests available with 2-(2H-benzotriazol-2-yl)-p-cresol

Reference	Type of study	Reliability (Klimisch)	Applied concentration	# of subjects evaluated	# (%) sensitized	comments
(Zhao and Li, 2014)	Patch test	2	1% in petrolatum	201	0 (0)	

Diagnostic patch tests

Patch tests in work place related sensitization patients (Stoeva et al., 2023)

The incidence of occupational contact dermatitis towards 2-(2H-benzotriazol-2-yl)-p-cresol was studied in dental staff with work-related dermatitis. Selected patients were patch tested for sensitization with 40 dental allergens in a cross-sectional study. The 329 subjects were dentists, dental staff and dental students who indicated to suffer from work-related dermatitis in a questionnaire. Of the participants, 73.9% were female and the mean age was 45.5 years (SD 13.5, age range 22-76 years). Clinical examination included documentation of contact dermatitis. Patch testing was performed with the Dental Screening Series (34 substances including 2-(2H-benzotriazol-2-yl)-p-cresol) as well as additional 6 allergens found in the dental work environment according to the European Society of Contact Dermatitis (ESCD) recommendations. Patches were applied to the upper back of the patients and removed after 48 hours. The sites were examined at removal (day 2), day 3/4 and after 7 days. While 32 of the 40 haptens tested produced positive reactions with 47.1% of participants showing at least one positive reaction, no positive responses were observed to 1% 2-(2H-benzotriazol-2-yl)-p-cresol. Therefore, no sensitization was observed in this group of workers with known dermatitis.

Patch tests in IVDK database (unpublished information (IVDK, 2024))

IVDK (Information Network of Departments of Dermatology, Germany) is a surveillance network of 56 German dermatological clinics monitoring contact allergy results to different substances (Schnuch et al., 2012). In this database, 202 patients (57 men, 145 women) were tested with 2-(2H-benzotriazol-2-yl)-p-cresol at 1% in Vaseline according to the AWMF guidance between 2002 and 2023 ("AWMF-S3-Leitlinie. Durchführung des Epikutantests mit Kontaktallergenen und Arzneimitteln," 2019). No positive reactions were observed (also no questionable reactions or irritations). No other concentrations were tested.

Summary human diagnostic patch tests

Table 11: Overview of clinical diagnostic patch tests available with 2-(2H-benzotriazol-2-yl)-p-cresol

Reference	Type of study	Reliability (Klimisch)	Applied concentration	# of subjects evaluated	# (%) sensitized	comments
(Stoeva et al., 2023)	Patch test	2	1% in petrolatum	329	0 (0)	Occupational dermatitis
(IVDK, 2024)	Patch test	2	1% in vaseline	202	0 (0)	Unpublished information
(Tarvainen, 1995)	Patch test	2	1% in petrolatum	343	0 (0)	
(Tomar et al., 2005)	Patch test	2	1% in petrolatum	50	1 (2)	
(Kanerva, et al., 1999)	Patch test	4	1% in petrolatum	357	0 (0)	

(Arisu et al., 1992)	Patch test	4	5% in petrolatum	30	0 (0)	
(Holness and Nethercott, 1997)	Patch test	4	2%	58	0	
(Peng et al., 2018)	Patch test	3	n.a.	88	7 (7.9)	No information on sensitization rate in non-facial contact dermatitis patients, no information on concentration tested

Human case studies

Few case reports have been published reporting allergic skin reactions to 2-(2H-benzotriazol-2-yl)-p-cresol in individual dermatitis patients due to its historical use in consumer products. Since the sensitization source (incl. concentration), contact intensity and habits are generally poorly described, these data can only be used as supporting data in a weight-of-evidence assessment.

Summary human case studies

Table 12: Overview of available case studies reporting positive patch tests to 2-(2H-benzotriazol-2-yl)-p-cresol in previously sensitized patients

Reference	Type of study	History	Reliability (Klimisch)	Patch result
(De Groot and Liem, 1983)	Case report	Edema and erythema of eyelid	4	Positive at 5%, questionable positive at 1%
(Hald et al., 2018)	Case report	History of eczema (3 occurrences) after repeated contact with plastic items	4	Positive reactions, concentrations tested not reported
(Cronin, 1980), reported in (Lee et al., 2019)	Case reports (secondary literature)	4 women with eczema on face using facial cream with substance	4	Positive at 1% in petrolatum
(Kullberg and Hylwa, 2020)	Case report	Genital dermatitis; use of sanitary pads	4	Positive at 1% in petrolatum

Reference	Type of study	History	Reliability (Klimisch)	Patch result
(Arisu et al., 1992)	Case report	History of cosmetic dermatitis, erythema after wearing new clothing	4	Positive at all concentrations tested
(Björkner and Niklasson, 1997)	Case report	Gingivitis in conjunction with dental filling	4	Positive to 1% in petrolatum
(Kaniwa et al., 1984)	Case report (abstract only)	Patient reacted to PUE tape used in T-shirt	4	No further information provided
(Niklasson and Björkner, 1989)	Case report	Eczema on palms, fingers and wrist due to watch strap	4	Positive to 1% in petrolatum
(Crépy et al., 2006)	Case report	Facial skin reactions to safety goggles	4	Positive to standard plastics battery of Chemotechnique® (1% in petrolatum)
(van Hecke and Vossaert, 1988)	Case report	Skin reaction to ostomy bag	4	Positive reaction to standard plastic and glues series of Chemotechnique® (1% in petrolatum)

Comparison with classification criteria:

Human data

According to CLP, Annex I, section 3.4.2.2.2.1, human evidence for sub-category 1A can include:

- (a) positive responses at $\leq 500 \mu\text{g}/\text{cm}^2$ (HRIPT, HMT - induction threshold);
- (b) diagnostic patch test data where there is a relatively high and substantial incidence of reactions in a defined population in relation to relatively low exposure;
- (c) other epidemiological evidence where there is a relatively high and substantial incidence of allergic contact dermatitis in relation to relatively low exposure.

According to CLP, Annex I, section 3.4.2.2.2.2, human evidence for sub-category 1B can include:

- (a) positive responses at $> 500 \mu\text{g}/\text{cm}^2$ (HRIPT, HMT – induction threshold);
- (b) diagnostic patch test data where there is a relatively low but substantial incidence of reactions in a defined population in relation to relatively high exposure;

- (c) other epidemiological evidence where there is a relatively low but substantial incidence of allergic contact dermatitis in relation to relatively high exposure.

The criteria for (a), (b) and (c) are discussed in further detail below and addressed again in the weight-of-evidence summary (Table 15, page 34 and 35).

- (a) Four human induction studies in healthy subjects are available (Table 9), one of which was assigned Klimisch score 2, while the other three tests were considered to be Klimisch 4 but can be used as supporting data. No skin sensitization was observed in any of the HRIPT and HMT available. From two studies, sufficient information was available for calculation of dose per skin area (DSA). In the Klimisch 2 HRIPT, a DSA of 590 $\mu\text{g}/\text{cm}^2$ was reached. In one of the Klimisch 4 studies, a DSA of 12916 $\mu\text{g}/\text{cm}^2$ could be calculated.

Therefore, the criteria for (a) show no skin sensitizing potential in human induction studies at DSA > 500 $\mu\text{g}/\text{cm}^2$. Thus, these data do support that any sensitization potential of 2-(2H-benzotriazol-2-yl)-p-cresol must be weak.

- (b) Several diagnostic patch test data are available on the substance under evaluation, both for professionals and patients with a history of sensitization (Table 11), as it is part of the standard testing battery for plastics and glues as well as the dental screening series, which are used in many countries. Four test series were identified with experimental description sufficient for assessment as reliable with restrictions (Klimisch 2), while three other publications were considered not assignable (Klimisch 4) due to lack of important information. One study was poorly reported in terms of methods and the results comprised only a subset of patients for which the selection criteria were not specified. Therefore, a reporting bias is suspected and a Klimisch score 3 (not reliable) was assigned. Among the Klimisch 2 publications, the sensitization rates were 0% in three studies, while one publication reported 1/50 subjects (2%) sensitized. However, it should be kept in mind that with only 50 subjects tested, a single positive reaction will already lead to sensitization rate of 2%, thus this value should be carefully weighed with the rest of the data. In the three Klimisch 4 studies, a total of 445 patient patch tests were reported, none of these patients showed a reaction to 2-(2H-benzotriazol-2-yl)-p-cresol. In the publication considered not reliable (Klimisch 3), 7 sensitized patients out of 88 (7.9%) were reported; however, only the results for a subset of the tested patients were reported for this substance (88 out of 443 patients tested in total); no information on sensitization prevalence among the whole group is provided. Based on the biased reporting and the lack of information on method and concentration tested, these data must be considered unreliable and be given less weight than the other data available.

Overall, applying the criteria suggested in ECHA's Guidance on the Application of the CLP Criteria (Table 13 and Table 14) (European Chemicals Agency, 2024), the majority of the studies available for (b) do support classification as Category 1B. One single study with relatively small patient numbers and insufficient detail of reporting provided a high frequency of sensitized dermatitis patients. However, the reliability of these results is questionable in the overall weight-of-evidence.

- (c) In addition to HRIPT and patch tests, there is one epidemiological study and several case reports with reported allergic reactions of single patients to 2-(2H-benzotriazol-2-yl)-p-cresol (Table 10, Table 12). Ten case studies were identified in the literature, reporting a total of 13 cases of sensitization to the substance under evaluation. While these data cannot be used for decision on sub-categorization, they are indicative of the relatively low sensitization frequency within the population, together with information from clinical patch testing.

Overall, the published and unpublished data available on human evidence shows a low/moderate frequency of sensitization when compared to the criteria as published in ECHA's Guidance on the Application of the CLP Criteria.

Table 13: Relatively high or low frequency of occurrence of skin sensitization (European Chemicals Agency, 2024)

Human diagnostic patch test data	High frequency	Low/moderate frequency	2-(2H-benzotriazol-2-yl)-p-cresol
General population studies	≥ 0.2 %	< 0.2 %	0% (0/201 in 1 study)
Dermatitis patients (unselected, consecutive)	≥ 1.0 %	< 1.0 %	n.a.
Selected dermatitis patients (aimed testing, usually special test series)	≥ 2.0 %	< 2.0 %	0.7 % (8/1128 in 7 studies)
Work place studies: 1: all or randomly selected workers 2: selected workers with known exposure or dermatitis	≥ 0.4 % ≥ 1.0 %	< 0.4 % < 1.0 %	n.a. 0% (0/329 in 1 study)
Number of published cases	≥ 100 cases	< 100 cases	13 (in 10 different publications)

Table 14: Relatively high or low exposure (European Chemicals Agency, 2024)

Exposure data	Relatively low exposure (weighting)	Relatively high exposure (weighting)	2-(2H-benzotriazol-2-yl)-p-cresol
Concentration/dose	< 1.0% < 500µg/cm ² (score 0)	≥ 1.0% ≥ 500µg/cm ² (score 2)	Score 2 (>500 µg/cm ² in HRIPT studies, 0.5 – 25 % in HRIPT studies)
Repeated exposure	< once/daily (score 1)	≥ once/daily (score 2)	Score 2 (Historically used in cosmetics including daily used products (CIR Expert Panel, 2008))
Number of exposures	< 100 exposures (score 0)	≥100 exposures (score 2)	Score 2 (Historically used in cosmetics including daily used products)

			(CIR Expert Panel, 2008))
--	--	--	---------------------------

Based on the human data available for evaluation, and applying the CLP criteria for skin sensitization, Category 1B is warranted.

Animal data

In the available GPMT, 80% and 90% of animals showed skin reactions at 24 and 48 hours after challenge, respectively. Intradermal induction was previously performed with 5% test substance. Some control animals also showed skin reactions, which is a well-known consequence of adjuvant use in intradermal induction. It would rather suggest that some reactions observed in the treated groups might be unspecific and thus present an overestimation of the response as compared to an underestimation. According to CLP Annex I, sections 3.4.2.2.3.2 and 3.4.2.2.3.3, a response rate of 80-90% after intradermal induction with 5% test substance would suggest a result eligible for sub-category 1B.

ECHA's Guidance on the Application of the CLP Criteria specifies that "Classification into sub-categories is required when data are sufficient (CLP Annex I 3.4.2.2.1.1). When Category 1A cannot be excluded, Category 1 should be applied instead of Category 1B. This is particularly important if only data are available from the guinea pig tests [...] showing a high response after exposure to a high concentration but where lower concentrations which could show the presence of such effects at lower doses are absent or in the absence of adequate dose-response information." (European Chemicals Agency, 2024)

From the data of the GPMT alone, classification as Category 1A cannot be excluded. However, data are available on a modified LLNA, which was conducted not in line with the OECD TG 429, but with the originally published protocol. Therefore, these data can be considered supporting evidence for decision on substance potency. Doses up to 2% were tested and no sensitizing properties of the substance were observed. While this result does not allow for a decision on skin sensitization in itself, it does show that in the LLNA a stimulation index <3 was observed at 2% and that consequently, an EC3 value would have to be >2%. According to CLP Annex I, sections 3.4.2.2.3.2 and 3.4.2.2.3.3, a LLNA EC3 value >2% is supportive of a classification as Category 1B.

Therefore, the available and relevant animal data indicate that sub-classification is possible and Category 1B is appropriate for classification of skin sensitization of 2-(2H-benzotriazol-2-yl)-p-cresol.

Summary

Based on the provisions in CLP Annex I, section 3.4.2.2.4, the following weight of evidence was derived:

Table 15: Weight of evidence consideration according to CLP Annex I, section 3.4.2.2.4

CLP Annex I, section 3.4.2.2.4	Data available	Results

<p>(a) positive data from patch testing, normally obtained in more than one dermatology clinic</p>	<p>Eight series of diagnostic patch tests are available (published and unpublished data). (For further information see Diagnostic patch tests and Table 11).</p>	<p>A total of 1457 patients (both professional workers and selected dermatitis patients) were described. Cases of positive patch tests were described in only 2/8 series. Incidences in these were 7/88 patients and 1/50 patients. Based on the reasons described above, the publication reporting 7/88 patients positive was considered unreliable and should be given less weight than other reports</p>
--	---	---

<p>(b) epidemiological studies showing allergic contact dermatitis caused by the substance. Situations in which a high proportion of those exposed exhibit characteristic symptoms are to be looked at with special concern, even if the number of cases is small</p>	<p>One patch test study in the general population which can be considered an epidemiological cross-sectional study is available. (For further information see section</p> <p>Epidemiological study</p> <p>Cross-sectional Patch test evaluation (Zhao and Li, 2014)).</p>	<p>201 university students were patch tested irrespective of their sensitization history. None of the subjects responded to a patch test with 2-(2H-benzotriazol-2-yl)-p-cresol.</p>
---	--	--

<p>(c) positive data from appropriate animal studies</p>	<p>One OECD TG 406 study (GLP) is available. Supporting information can be derived from a published skin sensitization test similar to OECD TG 429. (For further information, see sections Guinea pig maximization test (Ciba-Geigy Ltd., 1992) and Local Lymph Node Assay (Ikarashi et al., 1994a)).</p>	<p>In the OECD TG 406 study, 5% intradermal induction led to sensitization rates of 80% and 90% (24 and 48h, respectively). In the publication conducting an LLNA similar to OECD TG 429, no sensitization was observed at concentrations up to 2%.</p>
<p>(d) positive data from experimental studies in man</p>	<p>Data are available from three HRIPT and one HMT. (For further information see Human induction studies and Table 9)</p>	<p>One HRIPT was conducted with 0.5% test substance in 59 subjects while two other HRIPT were conducted with 25% test substance in 25 subjects each. In the HMT, 0.5% test substance was applied to 148 subjects. No skin sensitization was observed in any of these studies.</p>
<p>(e) well documented episodes of allergic contact dermatitis, normally obtained in more than one dermatology clinic</p>	<p>Ten publications with reports of case studies are available. (For further information see Table 12).</p>	<p>Nine reports are on single patients, one secondary publication reports on four cases. Generally, no information on sensitization induction is available, confirmatory patch testing was mostly positive at 1% elicitation concentration</p>
<p>(f) severity of reaction may also be considered</p>	<p>No reports of severe allergic reaction requiring hospitalization available</p>	<p>Most studies reporting sensitization in humans are reporting transient skin reactions without need for hospitalization; no information on sensitization induction and relevant exposure is available.</p>

As described above, both human and animal studies were considered to evaluate the sensitization potency of 2-(2H-benzotriazol-2-yl)-p-cresol. Animal data show a weak/moderate sensitization potential, which is further confirmed by a lack of sensitization in human induction studies and low frequencies of sensitization in human patch tests available. Epidemiological patch tests in the general population were negative and only few incidences of sensitization were reported in case studies. Overall, the available data are considered adequate and sufficient for sub-classification of 2-(2H-benzotriazol-2-yl)-p-cresol as skin sensitizer, Category 1B.

References

- Arisu, K., Hayakawa, R., Ogino, Y., Matsunaga, K., Kaniwa, M., 1992. Tinuvin® P in a spandex tape as a cause of clothing dermatitis. *Contact Dermatitis* 26, 311–316. <https://doi.org/10.1111/j.1600-0536.1992.tb00125.x>
- AWMF-S3-Leitlinie. Durchführung des Epikutantests mit Kontaktallergenen und Arzneimitteln., 2019.
- BASF SE (2011). EC 219-470-5: Daphnia magna Reproduction Test. Project No.: 51E0850/09E009. Report date: 2011-12-20.
- BASF SE (2018). 2-(2H-benzotriazol-2-yl)-p-cresol: Alga, Growth Inhibition Test with *Pseudokirchneriella subcapitata*, 72 hours. Project No.: 60E0850/09X200. Report date: 2018-03-09.
- BASF SE (2020). 14C-2-(2H-Benzotriazol-2'-yl)-4-methyl-phenol: Bioconcentration study in the rainbow trout (*Oncorhynchus mykiss*). Project no. 35F0235/19E005. Report date: 13 Aug 2020.
- Björkner, B., Niklasson, B., 1997. Contact allergy to the UV absorber Tinuvin P in a dental restorative material. *Am J Contact Dermat* 8, 6–7.
- Brooke D. and Crookes M. (2012). Depuration rate constant: growth correction and use as an indicator of bioaccumulation potential. Report LIT 7371.
- Ciba-Geigy Ltd. (1977a). Studies on the effects of UV-P on the hepatic xenobiotic metabolism enzymes in the rat. Report no: not available.
- Ciba-Geigy Ltd. (1977b). Distribution and elimination in the rat. Report date: 1977-11-30.
- Ciba-Geigy Ltd. (1988a). Report on Partition Coefficient. Report date: 1988-12-09.
- Ciba-Geigy Ltd. (1989). Report on the Test for Ready Biodegradability of TK 10047 in the Modified Strum Test (OECD-Guideline No. 301 B) (Paris 1981). Report no.: 88 44 64. Report date: 1989-02-21.
- Ciba-Geigy Ltd., 1992. Skin Sensitisation Test in the Guinea Pig Maximisation Test (study report No. 914119).
- CIR Expert Panel, 2008. Amended Final Report of the Safety Assessment of Drometrizole as used in Cosmetics1. *Int J Toxicol* 27, 63–75. <https://doi.org/10.1080/10915810802032412>
- Crépy, M.N., Grabas, A., Cohen Jonathan, A.M., Cuveillier, G., Choudat, D., 2006. Le Tinuvin P®, un nouvel allergène des lunettes de protection. *Archives des Maladies Professionnelles et de l'Environnement* 67, 754–755. [https://doi.org/10.1016/S1775-8785\(06\)70470-2](https://doi.org/10.1016/S1775-8785(06)70470-2)
- Cronin, E., 1980. Contact dermatitis.
- De Groot, A.C., Liem, D.H., 1983. Contact allergy to Tinuvin® P. *Contact Dermatitis* 9, 324–325. <https://doi.org/10.1111/j.1600-0536.1983.tb04410.x>
- European Chemicals Agency (ECHA; 202). Guidance on Information Requirements and Chemical Safety Assessment: Chapter R.11: PBT/vPvB assessment. Version 4.0. Publication date: December 2023. URL: <http://echa.europa.eu/web/guest/guidance-documents/guidance-on-reach> (accessed: 2024-03-18).

- European Chemicals Agency, 2024. Guidance on the application of the CLP criteria: guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures.
- Fischer C., J. Hiller, E. Leibold & T. Göen (2023). Toxikokinetik des UV-Absorbers 2-(2-Hydroxy-5-methylphenyl)-benzotriazol (UV-P) nach einmaliger oraler Applikation.
- Frankild, S., Basketter, D.A., Andersen, K.E., 1996. The value and limitations of rechallenge in the guinea pig maximization test. *Contact Dermatitis* 35, 135–140. <https://doi.org/10.1111/j.1600-0536.1996.tb02330.x>
- Gad, S.C., Dunn, B.J., Dobbs, D.W., Reilly, C., Walsh, R.D., 1986. Development and validation of an alternative dermal sensitization test: The mouse ear swelling test (MEST). *Toxicology and Applied Pharmacology* 84, 93–114. [https://doi.org/10.1016/0041-008X\(86\)90419-9](https://doi.org/10.1016/0041-008X(86)90419-9)
- Hald, M., Bergendorff, O., Isaksson, M., Johansen, J.D., 2018. Allergic contact dermatitis caused by plastic items containing the ultraviolet absorber drometrizole. *Contact Dermatitis* 79, 110–112. <https://doi.org/10.1111/cod.13007>
- Hill Top Research Institute, 1960. REPEATED INSULT PATCH TEST OF TINUVIN P IN DIMETHYL PHTHALATE SOLUTION (study report No. K-141C).
- Holness, D., Nethercott, J., 1997. Results of patch testing with a specialized collection of plastic and glue allergens. *American Journal of Contact Dermatitis* 8, 121–124. [https://doi.org/10.1016/S1046-199X\(97\)90006-3](https://doi.org/10.1016/S1046-199X(97)90006-3)
- Ikarashi, Y., Tsuchiya, T., Nakamura, A., 1993. A sensitive mouse lymph node assay with two application phases for detection of contact allergens. *Arch Toxicol* 67, 629–636. <https://doi.org/10.1007/BF01974070>
- Ikarashi, Y., Tsuchiya, T., Nakamura, A., 1994a. Contact sensitivity to Tinuvin® P in mice. *Contact Dermatitis* 30, 226–230. <https://doi.org/10.1111/j.1600-0536.1994.tb00649.x>
- Ikarashi, Y., Tsuchiya, T., Nakamura, A., 1994b. Contact sensitivity of and cross-sensitivity between 2-(2'-hydroxy-5'-methylphenyl)benzotriazole (Tinuvin® P) and 2-(2'-hydroxy-3'-tert-butyl-5'-methylphenyl)-5-chlorobenzotriazole (Tinuvin® 326) evaluated by lymph node cell proliferation and ear swelling response in mice. *Toxicology Letters* 71, 151–159. [https://doi.org/10.1016/0378-4274\(94\)90175-9](https://doi.org/10.1016/0378-4274(94)90175-9)
- IVDK, 2024. Data on 2-(2H-benzotriazol-2-yl)-p-cresol (CAS 2440-22-4) in IVDK database (unpublished information).
- Kanerva, L., Jolanki, R., Alanko, K., Estlander, T., 1999. Patch-test Reactions to Plastic and Glue Allergens. *Acta Dermato-Venereologica* 79, 296–300. <https://doi.org/10.1080/000155599750010706>
- Kaniwa, M., Kojima, S., Nakamura, A., Kanto, H., Ito, M., Ishihara, M., 1984. Chemical Approach to Contact Dermatitis Caused by Household Products. II. Analysis of Antioxidants in Commercial Rubber Gloves and Incidence of Positive Reactions to Antioxidants in Patch Testing. *Eisei Kagaku* 30, 126–137. https://doi.org/10.1248/jhs1956.30.3_126
- Kimber, I., Weisenberger, C., 1989. A murine local lymph node assay for the identification of contact allergens: Assay development and results of an initial validation study. *Arch Toxicol* 63, 274–282. <https://doi.org/10.1007/BF00278640>
- Kligman, A.M., 1964a. Sensitization, Irritation and Phototoxicity (summary report).

- Kligman, A.M., 1964b. Repeated Insult Patch Test (summary report).
- Kligman, A.M., Basketter, D.A., 1995. A critical commentary and updating of the guinea pig maximization test. *Contact Dermatitis* 32, 129–134. <https://doi.org/10.1111/j.1600-0536.1995.tb00801.x>
- Kullberg, S.A., Hylwa, S.A., 2020. The versatile UV absorber drometrizole: Expanding from the realm of cosmetics to feminine sanitary products. *Contact Dermatitis* 83, 518–519. <https://doi.org/10.1111/cod.13651>
- Lee, J.K., Kim, K.-B., Lee, J.D., Shin, C.Y., Kwack, S.J., Lee, B.-M., Lee, J.Y., 2019. Risk Assessment of Drometrizole, a Cosmetic Ingredient used as an Ultraviolet Light Absorber. *ToxicolRes* 35, 119–129. <https://doi.org/10.5487/TR.2019.35.2.119>
- Magnusson, B., 1980. Identification of contact sensitizers by animal assay. *Contact Dermatitis* 6, 46–50. <https://doi.org/10.1111/j.1600-0536.1980.tb03894.x>
- Niklasson, B., Björkner, B., 1989. Contact allergy to the UV-absorber Tinuvin P in plastics. *Contact Dermatitis* 21, 330–334. <https://doi.org/10.1111/j.1600-0536.1989.tb04753.x>
- NITE (1996). Ready biodegradability study of 2-(2'-hydroxy-5'-methylphenyl) benzotriazole (test substance No. K-1341). Study no.: 21341. Report date: 1996-10-23.
- NITE (1998). Concentration study of 2-(2'-hydroxy-5'-methylphenyl) benzotriazole (test substance No. K-1341) in carp. Study no.: 51341. Report date: 1998-03-26.
- OECD (2011). Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test. OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, adopted: 23 March 2006, Annex 5 corrected: 28 July 2011. <https://doi.org/10.1787/9789264069923-en>.
- OECD (2012). Test No. 305: Bioaccumulation in Fish: Aqueous and Dietary Exposure, OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, adopted 02 October 2012, <https://doi.org/10.1787/9789264185296-en>.
- OECD (2017). Guidance Document on Aspects of OECD TG 305 on Fish Bioaccumulation. Series on Testing & Assessment No. 264. ENV/JM/MONO(2017)16.
- OECD (2019). Guidance Document on Aqueous-Phase Aquatic Toxicity Testing of Difficult Test Chemicals. Series on Testing and Assessment No. 23 (Second Edition). ENV/JM/MONO(2000)6/REV1.
- Peng, F., Mu, Z., He, C., Xue, C., Li, W., Wang, Q., Chen, Z., Zhang, J., 2018. Patch testing in facial dermatitis using Chinese Baseline Series (60 allergens) and Cosmetic Series (58 allergens). *Acad Dermatol Venereol* 32. <https://doi.org/10.1111/jdv.14822>
- Schnuch, A., Geier, J., Lessmann, H., Arnold, R., Uter, W., 2012. Surveillance of contact allergies: methods and results of the Information Network of Departments of Dermatology (IVDK). *Allergy* 67, 847–857. <https://doi.org/10.1111/j.1398-9995.2012.02834.x>
- Sijm D.T.H.M., Verberne M.E., de Jonge W.J., Pärt P. and Opperhuizen A. (1995). Allometry in the uptake of hydrophobic chemicals determined in vivo and in isolated perfused gills. *Toxicol. Appl. Pharmacol.* 131: 130-135.
- Stoeva, I., Kanazirska, P., Dogandzhiyska, V., Karova, E., Tomova, Z., Chervenkov, L., 2023. Patch test results in dental staff with dermatitis. *Aust J Dermatology* 64, 346–353. <https://doi.org/10.1111/ajd.14112>

- Tarvainen, K., 1995. Analysis of patients with allergic patch test reactions to a plastics and glues series. *Contact Dermatitis* 32, 346–351. <https://doi.org/10.1111/j.1600-0536.1995.tb00623.x>
- Tomar, J., Jain, V.K., Aggarwal, K., Dayal, S., Gupta, S., 2005. Contact Allergies to Cosmetics: Testing with 52 Cosmetic Ingredients and Personal Products. *The Journal of Dermatology* 32, 951–955. <https://doi.org/10.1111/j.1346-8138.2005.tb00880.x>
- van Hecke, E., Vossaert, K., 1988. Allergic contact dermatitis from an ostomy bag. *Contact Dermatitis* 18, 121–122. <https://doi.org/10.1111/j.1600-0536.1988.tb02764.x>
- Yamano, T., Noda, T., Shimizu, M., Morita, S., 1993. Allergenicity evaluation of chemicals for use in household products, Contact allergenicity of tinuvin P and tinuvin 326 in guinea pigs. *Annual Report of Osaka City Institute of Public Health and Environmental Sciences* 55, 47–52.
- Yamano, T., Shimizu, M., Noda, T., 2001. Relative Elicitation Potencies of Seven Chemical Allergens in the Guinea Pig Maximization Test. *JOURNAL OF HEALTH SCIENCE* 47, 123–128. <https://doi.org/10.1248/jhs.47.123>
- Zhang, Shuying; Zhongyu Wang; Jingwen Chen; Qing Xie; Minghua Zhu; Wenjing Han (2021). Tissue-Specific Accumulation, Biotransformation, and Physiologically Based Toxicokinetic Modeling of Benzotriazole Ultraviolet Stabilizers in Zebrafish (*Danio rerio*). *Environ. Sci. Technol.* 2021, 55, 11874–11884.
- Zhao, J., Li, L., 2014. Contact sensitization to cosmetic series of allergens in a general population in Beijing. *J of Cosmetic Dermatology* 13, 68–71. <https://doi.org/10.1111/jocd.12078>

Appendix I: *Daphnia magna* reproduction study

Table 16: Analytically measured concentrations of 2(2-H-benzotriazol-2-yl)-p-cresol

Nominal Concentration [µg/L]	Renewal period	Analytically measured concentration				
		Initial [µg/L]	48-72h old [µg/L]	Time-weighted Mean		
				[µg/L] ^a	[%] ^b	[%] ^c
0 (control)	Day 2 - 4	0	0	0	–	–
	Day 11 - 14	0	0			
	Day 16 - 18	0	0			
SC	Day 2 - 4	0	0	0	–	–
	Day 11 - 14	0	0			
	Day 16 - 18	0	0			
1.3	Day 2 - 4	1.3	0.7	0.72	56%	56%
	Day 11 - 14	1.3	0.2			
	Day 16 - 18	1.3	0.3			
4.1	Day 2 - 4	4.6	2.7	2.5	61%	56%
	Day 11 - 14	4.5	0.6			
	Day 16 - 18	4.3	1.0			
13.0	Day 2 - 4	14.2	7.8	8.3	64%	57%
	Day 11 - 14	15.4	2.6			
	Day 16 - 18	14	3.5			
41	Day 2 - 4	38.6	23.5	21.6	53%	47%
	Day 11 - 14	46.2	18.3			
	Day 16 - 18	54.0	-			
130	Day 2 - 4	133	109	120.4	93%	83%
	Day 11 - 14	[142]	–			
	Day 16 - 18	[162]	–			

a: based on 7 renewal period days;
b: % of Nominal concentration;
c: % of Mean Initial
Numbers in bold are the mean of 2 retained samples.

Table 17: Overview on algal concentration (green coloration) in different cultures with *Daphnia magna* over a period of 72 hours: 1) Young daphnids, 2) Adult daphnids

	Incubation time			
	0 h	24 h	48 h	72 h
Beakers containing young daphnids in M4 medium fed with 0.10 mg TOC algae/day				
Young daphnid before feeding				
Young daphnid right after feeding				
M4 medium without daphnid, only containing 0.10 mg TOC algae/day				
Beakers containing adult daphnids in M4 medium fed with 0.20 mg TOC algae/day				
Adult daphnid before feeding				
Adult daphnid right after feeding				
M4 medium without daphnid, only containing 0.20 mg TOC algae/day				