SI1 - Supplementary information to:

Managing PMT/vPvM Substances in Consumer Products through the Concepts of Essential Use and Functional Substitution: a Case-Study for Cosmetics

Joanke van Dijk[1, 2,]*, Romain Figuiere[3], Stefan Dekker[1], Annemarie van Wezel[2], Ian Cousins[3]

1 Copernicus Institute of Sustainable Development, Utrecht University, 3584 CB Utrecht, the Netherlands; 2 Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, PO Box 94240, 1090 GE Amsterdam, the Netherlands; 3 Department of Environmental Science, Stockholm University, SE-10691 Stockholm, Sweden.

*Corresponding author j.vandijk3@uu.nl

Table of Contents

This SI document provides more detailed information on the methods used in the study and an overview of the type of data contained in the other SI excel files.

SI2 – List of PMT/vPvM substances contained in cosmetic products

The table S2 is listing the substances from the Cosing database which have been identified as PMT or vPvM according to the analysis from the German Environmental Agency (Arp and Hale, 2019) along with their chemical functions, and the type of cosmetic products they are sued based on the information contained in Kemiluppen and Cosmethics databases.

Table S2 – List of PMT/vPvM substances listed in Cosing database, their functions and the type of products in which they are used (see separate Excel file SI2)

SI3 - Shortlisting of alternatives for further assessment

The figure 1 below presents the workflow which has been followed to shortlist chemical alternatives for further hazard assessment. In short, the Annex III of the Cosmetic Products Regulation (CPR) lists the substances subject to restriction of use under the CPR; SUBSPORT portal is a database which gathers 32 lists from authorities, industry or NGOs of substances for which the uses are restricted either legally or voluntarily (the lists available are compiled in Table SI3.1); the PACT tool is a tool provided by ECHA which summarize risk assessment and risk management activities in the EU; and the REACH Annex III Inventory (Link) compiles a list of substances which are likely to meet the criteria of Annex III of REACH based on QSAR data.

The following data is included in SI3:

- A full table of lists included in the SUBSPORT portal:
    - o **Table S3.1: list of restricted substances from authorities, industry or non-governmental organisations (NGOs) included in the SUBSPORT portal**
- All the information which has been gathered to identify shortlisted candidates for further hazard assessment:
    - o **Table S2.2.1: Authoritative list screening of potential alternatives to the substances of interest**
- Shortlisted chemical alternative which have been selected for further hazard assessment, along with their CAS number, molecular structure and physico-chemical properties:
    - o **Table S2.2.2: List of shorlisted potential alternatives for further assessment (see separated Excel file SI2)**
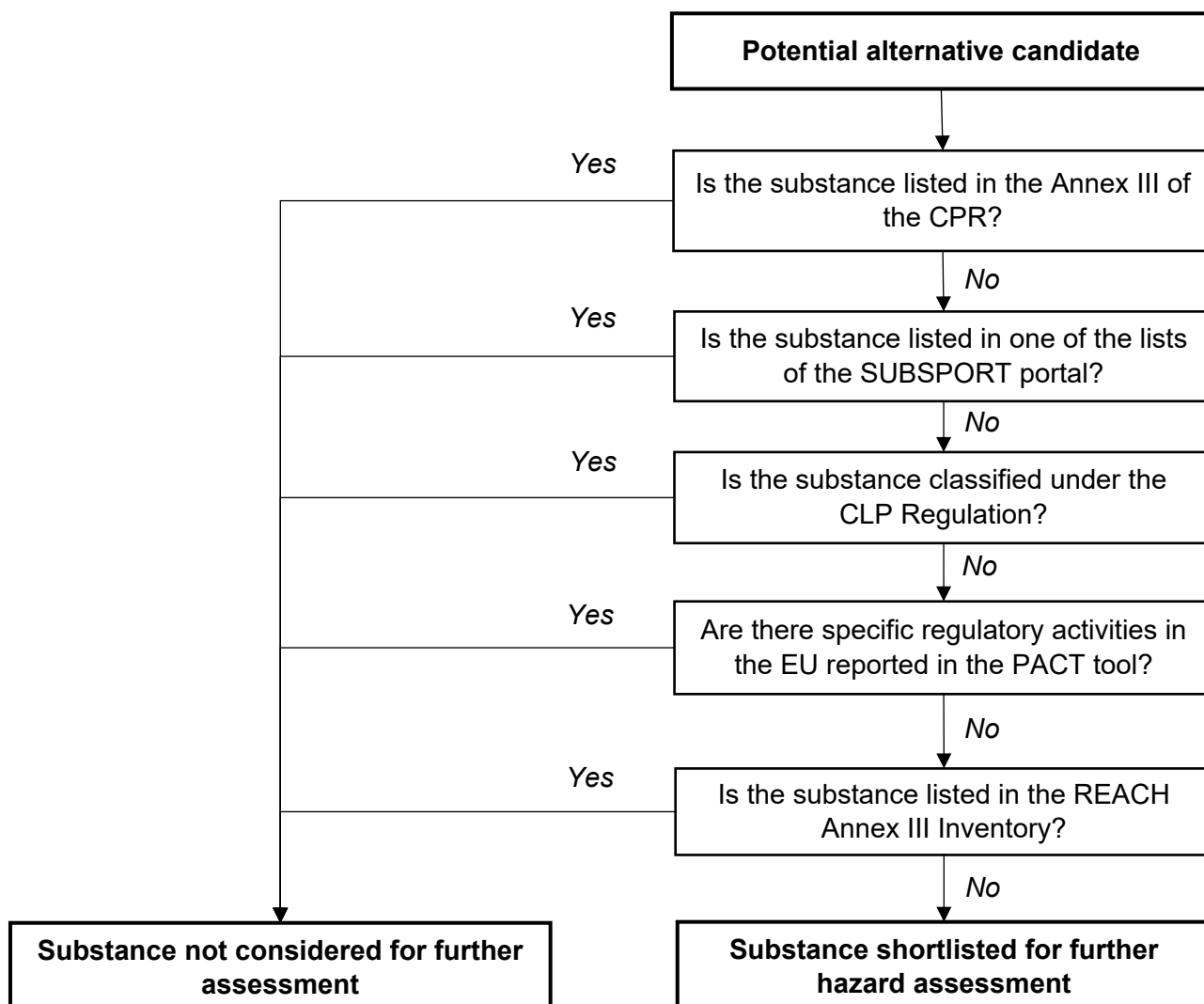
Figure 1 Workflow for selecting shortlisted chemical alternative for further hazard assessment: Annex III of the Cosmetic Products Regulation (CPR) lists the substances subject to restriction of use under the CPR; SUBSPORT portal is a database which gathers 32 lists from authorities, industry or NGOs of substances for which the uses are restricted either legally or voluntarily (the lists available are compiled in Table SI3.1); the PACT tool is a tool provided by ECHA which summarize risk assessment and risk management activities in the EU; and the REACH Annex III Inventory (Link) compiles a list of substances which are likely to meet the criteria of Annex III of REACH based on QSAR data.

SI4 – Hazard profiling of the shortlisted alternatives

In SI4, all the hazard data which have been collected for all hazard endpoints and all shortlisted chemical alternatives are presented. More detailed information on the considered endpoints and how data was collected is given below.

1. Hazard endpoints considered

The Table 1 below is listing the hazard endpoints which have been considered to perform the hazard assessment of the shortlisted alternatives.

Table 1 Hazard endpoints considered for the hazard assessment. P = Persistency; B = Bioaccumulation potential; M = Mobility; T human = Toxicity to human health; T env = Toxicity to the environment

Parameter	Endpoint considered
P	Ready biodegradability
	Anaerobic degradation
	Half-life soil
	Half-life sediment
	Half-life water
	Half-life air
B	Octanol-water partition coefficient
	Bioconcentration factor
M	Organic carbon-water partition coefficient
T human	Carcinogenicity
	Mutagenicity
	Teratogenicity
	Acute toxicity – Oral
	Acute toxicity – Dermal*
	Skin sensitisation
	Skin irritation
	Eye irritation
	Specific Target Organ Toxicity – Repeated Exposure – Oral
	Endocrine disruption: a) Androgenic b) Estrogenic c) Thyroid
T env	Acute toxicity – Algae
	Acute toxicity – Daphnia
	Acute toxicity – Fish
	Chronic toxicity – Daphnia
	Chronic toxicity – Daphnia

*: Acute toxicity via dermal exposure has been considered only for the assessment of alternatives to benzophenone-4 as no data was available for this endpoint for the other alternatives.

2. Gathering of hazard data

Experimental hazard data were collected from the sources below (organised by order of preferences):

i. ECHA Registration database (https://echa.europa.eu/information-on-chemicals/registered-substances)

ii. ComTox database (https://comptox.epa.gov/dashboard/)

iii. EcoTox database (https://cfpub.epa.gov/ecotox/)

iv. OECD eChemPortal (https://www.echemportal.org/echemportal/)

v. EFSA publications (https://www.efsa.europa.eu/en)

vi. Open scientific literature. The terms used for the open literature search are listed per endpoints in Table 2

Table 2 Terms for open scientific literature search for every hazard endpoint considered

Parameter	Endpoint considered	Search terms
P	Ready biodegradability	"biodegradation"
	Anaerobic degradation	NA
	Half-life soil	"half life" "soil"
	Half-life sediment	"half life" "sediments"
	Half-life water	"half life" "water"
	Half-life air	"half life" "air"
B	Octanol-water partition coefficient	"Kow"
	Bioconcentration factor	"BCF"
M	Organic carbon-water partition coefficient	"Koc"
T human	Carcinogenicity	"Carcinogenic*" "in vitro"
	Mutagenicity	"Mutagenic*" "in vitro"
	Teratogenicity	"reproduction" "toxicity" "in vitro" "reprotoxic"
	Acute toxicity – Oral	"toxicity" "oral" "rats" "mice" "roddens" "in vitro"
	Acute toxicity – Dermal*	"toxicity" "dermal" "rats" "mice" "roddens" "in vitro"
	Skin sensitisation	NA
	Skin irritation	NA
	Eye irritation	NA
	Specific Target Organ Toxicity – Repeated Exposure – Oral	"organ" "chronic" "toxicity" "rats" "mice" "rodent" "in vitro" "mammalian"
	Endocrine disruption: d) Androgenic e) Estrogenic f) Thyroid	"endocrine*"
T env	Acute toxicity – Algae	"green algae" "toxicity"
	Acute toxicity – Daphnia	"daphnia magna" "toxicity" "invertebrate"
	Acute toxicity – Fish	"Danio rerio" "Toxicity" "zebrafish" "rainbow trout" "Pimaphales promelas" "in vitro" "embryo test"
	Chronic toxicity – Daphnia	"daphnia magna" "toxicity" "invertebrate"
	Chronic toxicity – Daphnia	"Danio rerio" "Toxicity" "zebrafish" "rainbow trout"

If different sources provided different effect concentrations for a same hazard endpoint, the lowest concentration was kept for the assessment.

If experimental data was not available, QSAR models available in the OECD QSAR Toolbox (v 4.5) and the VEGA platform were used to fill in further data gaps according to the ECHA guidance on how to use QSAR models (ECHA, 2016). According to the point 1.3 of the Annex VI to Regulation (EC) No 1907/2006, results of QSARs may be used instead of testing when the following conditions are met:

1. results are derived from a QSAR model whose scientific validity has been established;
2. the substance falls within the applicability domain of the QSAR model;
3. results are adequate for the purpose of classification and labelling and/or risk assessment, and;
4. adequate and reliable documentation of the applied method is provided.

It has been assumed that 1st, 3rd and 4th points are automatically met for models available in the OECD QSAR toolbox and the VEGA platform. The Table 3 below lists the number of models available for each hazard endpoints considered in this study. QSAR predictions from a specific model were used only if the substance was within the applicability domain of the model.

Table 3 Number of QSAR models available for each hazard endpoint

Parameter	Assessment criteria	Number of models	Software platforms
P	Ready biodegradability	6	OECD QSAR Toolbox & VEGA
	Half-life soil	1	OECD QSAR Toolbox
	Half-life sediment	1	OECD QSAR Toolbox
	Half-life water	1	OECD QSAR Toolbox
	Half-life air	1	OECD QSAR Toolbox
	Anaerobic degradation	1	OECD QSAR Toolbox
B	LogKow	4	OECD QSAR Toolbox & VEGA
	BCF	11	OECD QSAR Toolbox & VEGA
M	LogKoc	3	OECD QSAR Toolbox & VEGA
T human	Carcinogenicity	23	OECD QSAR Toolbox & VEGA
	Mutagenicity	74	OECD QSAR Toolbox & VEGA
	Reprotoxicity (teratogenicity)	7	OECD QSAR Toolbox & VEGA
	Acute toxicity (oral, dermal exposure)	3	OECD QSAR Toolbox & VEGA
	Skin sensitisation	8	OECD QSAR Toolbox & VEGA
	Skin irritation	5	OECD QSAR Toolbox
	Eye irritation	1	OECD QSAR Toolbox
	Specific Target Organ Toxicity - Repeated Exposure (oral)	4	OECD QSAR Toolbox
	Endocrine disruption: *Androgenic* *Estrogenic* *Thyroid*	29 *5* *14* *10*	OECD QSAR Toolbox & VEGA
T env	Algae (EC50)	6	OECD QSAR Toolbox & VEGA
	Daphnids acute	8	OECD QSAR Toolbox & VEGA
	Fish acute	14	OECD QSAR Toolbox & VEGA

Daphnids chronic	2	OECD QSAR Toolbox & VEGA	
Fish chronic	2	OECD QSAR Toolbox & VEGA	

If several QSAR models were available for a quantitative hazard endpoint (e.g. acute toxicity to algae), the average of the predictions from models for which the substance was within the applicability domain was kept for further assessment. In case of qualitative hazard endpoint (e.g. carcinogenicity), a hazard endpoint score was determined for a specific substance by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

The equation illustrates the calculation of the mutagenicity score for Bornelone (CAS: 2226-11-1).

$$Mutagenicity\ score_{Bornelone} = \frac{Number\ of\ models\ predicting\ a\ negative\ outcome}{Number\ of\ models\ for\ which\ Bornelone\ is\ in\ the\ AD} = \frac{55}{58} = 0.94$$

SI5 – Hazard assessment and comparison of the alternatives

1. Harmonization of the hazard data

For some hazard endpoints, the experimental results and/or QSAR predictions were documented as qualitative data. In order to compare hazard profiles of alternatives it was necessary to transform these qualitative results into a quantitative score. Furthermore, for some endpoints, it was necessary to log-transformed the quantitative results so the distribution of the data would be closer to a normal distribution.

This part is summarizing all the changes which have been made to the data for each hazard endpoint considered.

- *Ready biodegradability*

Experimental data were documented qualitatively. The results were transformed into quantitative data by assigning a score equalled to 0 if the substance is not readily biodegradable or 1 if the substance is readily biodegradable.

If experimental data were not available, a combined prediction from BIOWIN3 and BIOWIN5 models was used. The results were transformed in similar manner as for experimental data.

- *Half-life in soil*

No transformation of the data was made for this endpoint.

- *Half-life in sediments*

No transformation of the data was made for this endpoint.

- *Half-life in water*

No transformation of the data was made for this endpoint.

- *Half-life in air*

No transformation of the data was made for this endpoint.

- *Anaerobic biodegradation*

No transformation of the data was made for this endpoint.

- *Octanol-water partition coefficient (Kow)*

The data on octanol-water partition coefficient were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Bioconcentration factor (BCF)*

The data on the bioconcentration factor were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Octanol-water partition coefficient (Koc)*

The data on organic carbon-water partition coefficient were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Carcinogenicity*

Qualitative results were transformed into quantitative data by assigning a score equalled to 0 if the substance is carcinogenic or 1 if the substance is not carcinogenic. In some cases, the experimental results were expressed quantitatively. In those cases, the final conclusion of the study on carcinogenicity potential was documented and transformed as explained above.

The predictions from QSAR models for carcinogenicity were qualitative. The carcinogenicity score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Mutagenicity*

Qualitative results were transformed into quantitative data by assigning a score equalled to 0 if the substance is mutagenic or 1 if the substance is not mutagenic.

The predictions from QSAR models for mutagenicity were qualitative. The mutagenicity score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Teratogenicity*

The predictions from QSAR models for teratogenicity were qualitative. The teratogenicity score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

The experimental data to evaluate the teratogenicity of a specific substance were quantitative. In those cases, the final qualitative conclusion of the study on the teratogenicity potential of the substance was documented. A score of 0 was assigned to the substance if it was concluded that it presented a teratogenic potential, and a score of 1 if it was concluded that the substance is not teratogenic.

- *Acute toxicity – Oral exposure*

The data on acute toxicity were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Acute toxicity – Dermal exposure*

The data on acute toxicity were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Skin sensitization*

Qualitative results were transformed into quantitative data by assigning a score equalled to 0 if the substance is skin sensitizing or 1 if the substance is not skin sensitizing.

The predictions from QSAR models for skin sensitizing were qualitative. The skin sensitizing score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Skin irritation*

Qualitative results were transformed into quantitative data by assigning a score equalled to 0 if the substance is skin irritating or 1 if the substance is not skin irritating.

The predictions from QSAR models for skin sensitizing were qualitative. The skin irritation score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Eyes irritation*

Qualitative results were transformed into quantitative data by assigning a score equalled to 0 if the substance is eyes irritating or 1 if the substance is not eyes irritating. If the data suggested that the substance was possibly irritating to the eyes, a score of 0.5 was assigned.

The predictions from QSAR models for skin sensitizing were qualitative. The skin irritation score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Specific Targeted Organ Toxicity – Repetitive Exposure (STOT RE)*

The predictions from QSAR models for STOT-RE were qualitative. The STOT-RE score of a specific substance was then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

Based on the information provided on the development of the model available, the model returns a "positive" result if the predicted effect concentration is within the range 0.0167-2.69 mg/kg bw/day, and a "negative" result if the predicted concentration is within the range 5-1000 mg/kg bw/day. These categories were used to expressed the experimental results qualitatively.

- *Endocrine disruption potential – Androgen, Estrogen or Thyroid*

The predictions from QSAR models for endocrine disruption potential were qualitative. The endocrine disruption to androgen, estrogen and thyroid activities score of a specific substance were then calculated by calculating the ratio of the number of models returning a negative outcome over the total number of model available for which the substance was within the applicability domain.

- *Acute toxicity to algae*

The data on acute toxicity to algae were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Acute toxicity to daphnia*

The data on acute toxicity to daphnia were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Acute toxicity to fish*

The data on acute toxicity to fish were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Chronic toxicity to daphnia*

The data on acute toxicity to daphnia were log-transformed in order to have a distribution of the data closer to a normal distribution.

- *Chronic toxicity to fish*

The data on acute toxicity to fish were log-transformed in order to have a distribution of the data closer to a normal distribution.

2. Normalisation of the data and handling of data gaps

2.1. Normalisation of the hazard data

In order to compare the hazard profiles of the chemical alternatives with certain multi-criteria decision analysis (MCDA) methods, it was necessary to normalise the hazard data. Therefore, for a specific endpoint, the hazard data were transformed so it is within the range $0 - 1$, with 0 representing the worst level, and 1 representing the best level, as it has been done in previous studies (Zheng et al., 2019). Table 4 below specifies to what 0 and 1 in normalised data correspond to in the original data, for each hazard endpoint.

Table 4 Levels for each hazard endpoint for normalisation of the data

Parameter	Endpoint	Level 0	Level 1
P	Ready biodegradability	0[b]	1[b]
	Half-life soil	Maximum among alternatives[a]	0[a]
	Half-life sediment	Maximum among alternatives[a]	0[a]
	Half-life water	Maximum among alternatives[a]	0[a]
	Half-life air	Maximum among alternatives[a]	0[a]
	Anaerobic degradation	Minimum among alternatives[c]	1.5[c]
B	LogKow	Maximum among alternatives[d]	-3[d]
	BCF	Maximum among alternatives[a]	-1[a]
M	LogKoc	-2[a]	Maximum among alternatives[a]
T human	Carcinogenicity	0[b]	1[b]
	Mutagenicity	0[b]	1[b]
	Reprotoxicity (teratogenicity)	0[b]	1[b]
	Acute toxicity (oral, dermal exposure)	1[a]	Maximum among alternatives[a]
	Skin sensitisation	0[b]	1[b]
	Skin irritation	0[b]	1[b]
	Eye irritation	0[b]	1[b]
	Specific Target Organ Toxicity -Repeated Exposure (oral)	0[b]	1[b]
	Endocrine disruption:		1[b]

	Androgenic	0 [b]	1 [b]
	Estrogenic	0 [b]	1 [b]
	Thyroid	0 [b]	
T env	Algae (EC50)	1x10^{-6} [a]	Maximum among alternatives [a]
	Daphnids acute	1x10^{-6} [a]	Maximum among alternatives [a]
	Fish acute	1x10^{-6} [a]	Maximum among alternatives [a]
	Daphnids chronic	1x10^{-6}	Maximum among alternatives
	Fish chronic	1x10^{-6}	Maximum among alternatives

a: Same approach as in Zheng et al., 2019; b: for hazard endpoint for which results were expressed qualitatively, the hazard endpoints scores were already normalized, so no transformation was needed; c: for anaerobic degradation, the "best level" was considered to 1.5 as it is the maximum prediction from the QSAR model which was used; d: for log(Kow), it has been considered that a high log(Kow) would represent a greater concern, and the "best level" was assigned arbitrarily to ensure that all the data would be included within the range.

2.2.   Data gaps and sensitivity analysis

As it has been done by to Zheng et al. (2019), three different scenarios were considered in order to evaluate the sensitivity of the MCDA models to data gaps. In each scenario, data gaps were replaced by a different value in the normalised data (Zheng et al., 2019):

- A risk neutral scenario, in which the data gaps were replaced by the value 0.5;
- A risk averse scenario, in which data gaps were replaced by the value 0.8;
- A risk seeking scenario, in which data gaps were replaced by the value 0.2.

3.   Multicriteria decision analysis methods

3.1. Heatmap

For the heat map, the thresholds of the different categories were determined according to legislation levels (i.e. CLP Regulation) or literature, as suggested by Zheng et al. (2019). For qualitative hazard endpoint which were not considered by Zheng et al. (2019) (e.g. endocrine disruption potential), a similar approach as for other qualitative hazard endpoint (e.g. carcinogenicity) was taken to determine the thresholds. For chronic toxicity on daphnia and fish endpoints, the same threshold as for acute toxicity endpoints were taken. For log(Kow), the threshold were determined based on a report from ECETOC which is linking the bioconcentration factor with the value of log(Kow). Hence, the threshold of log(BCF) were used to determine the threshold for log(Kow) (ECETOC, 2000). For anaerobic degradation, values from the BIOWIN user guide were taken to set thresholds (US EPA, 2012). Table 5 below presents the thresholds of each categories for every hazard endpoints considered.

Table 5 Threshold for each colour category to create the heat map

Parameter	Endpoint	RED	ORANGE	YELLOW	GREEN
P	Ready biodegradability[b]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Half-life soil (days)[a]	>180	120 – 180	60 – 120	<60
	Half-life sediment (days) [a]	>180	120 – 180	60 – 120	<60
	Half-life water (days) [a]	>180	120 – 180	60 – 120	<60
	Half-life air (days) [a]	>180	120 – 180	60 – 120	<60
	Anaerobic degradation[b]	<-0.5	-0.5 – 0	0 – 0.5	>0.5
B	LogKow[b]	[5.71;781]	[5.20;7.71] [7.81;8.10]	[4.41;5.20] [8.10;8.54]	logKow < 4.41 logKow > 8.54
	LogBCF [a]	>3.70	3.30 – 3.70	2.70 – 3.30	<2.70
M	LogKoc [a]	<4.5	4.5 – 6.5	6.5 – 9	>9
T human	Carcinogenicity [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Mutagenicity [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Reprotoxicity (teratogenicity) [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Acute toxicity (oral, dermal) (mg/kg bw) [a]	<50	300 – 50	300 – 2000	>2000
	Skin sensitisation [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Skin irritation [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Eye irritation [a]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Specific Target Organ Toxicity -Repeated Exposure (oral) [b]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
	Endocrine disruption: *Androgenic* [b] *Estrogenic* [b] *Thyroid* [b]	0 – 0.2	0.2 – 0.5	0.5 – 0.8	0.8 – 1
T env	Algae (EC50) [a]	<1	1 – 10	10 – 100	>100
	Daphnids acute [a]	<1	1 – 10	10 – 100	>100
	Fish acute [a]	<1	1 – 10	10 – 100	>100
	Daphnids chronic[b]	<1	1 – 10	10 – 100	>100
	Fish chronic[b]	<1	1 – 10	10 – 100	>100

a: The threshold was taken from Zheng et al. (2019); b: See explanations above

A hazard score was calculated for each alternative by assigning a value to the colour categories in the heat map: red category was scored 4; orange category was scored 3; yellow category was scored 2; green category was scored 1. The final hazard score of an alternative was obtained by summing up the scores of every hazard endpoint considered. Data gaps were scored 2.5, 3.4 and 1.6 in the risk neutral, risk averse and risk seeking scenarios, respectively.

The resulting heatmaps and the total hazard scores for all alternatives for each case study chemicals and for every scenarios (i.e. risk neutral, risk averse and risk seeking) are presented in table S5.1.

Table S5.1: Heatmap of chemical alternatives for each case study chemical for risk neutral, risk averse and risk seeking scenarios

3.2. Multi-attribute utility theory (MAUT)

In the MAUT approach, partial scores for persistency (P), bioaccumulation potential (B), mobility (M), toxicity to human health and toxicity to the environment (T) were determined by calculating the average value normalised data of the corresponding hazard endpoints. In this study it was assumed that every hazard endpoint is equally important and therefore have the same weight in the decision-making. Hence, the final PBMT scores of each alternative was determined by summing up all partial scores for P, B, M and T. Potential alternatives were then compared between each other and with the substance to phase out based on their final PBMT score. It was assumed that the alternative with a higher PBMT score could be considered as safer. Figure 2 below is presenting the final PBMT scores for every alternative to each case study chemicals in the risk neutral, risk averse and risk seeking scenarios.

The data used for the MAUT approach are presented in tables S5.2A, 2B and 2C.

Table S5.2A: Data used in the MAUT approach to identify safer alternatives to Allura Red

Table S5.2A: Data used in the MAUT approach to identify safer alternatives to Benzophenone-4

Table S5.2A: Data used in the MAUT approach to identify safer alternatives to Climbazole

Data for the risk neutral, risk averse and risk seeking scenarios is presented on the same sheet.
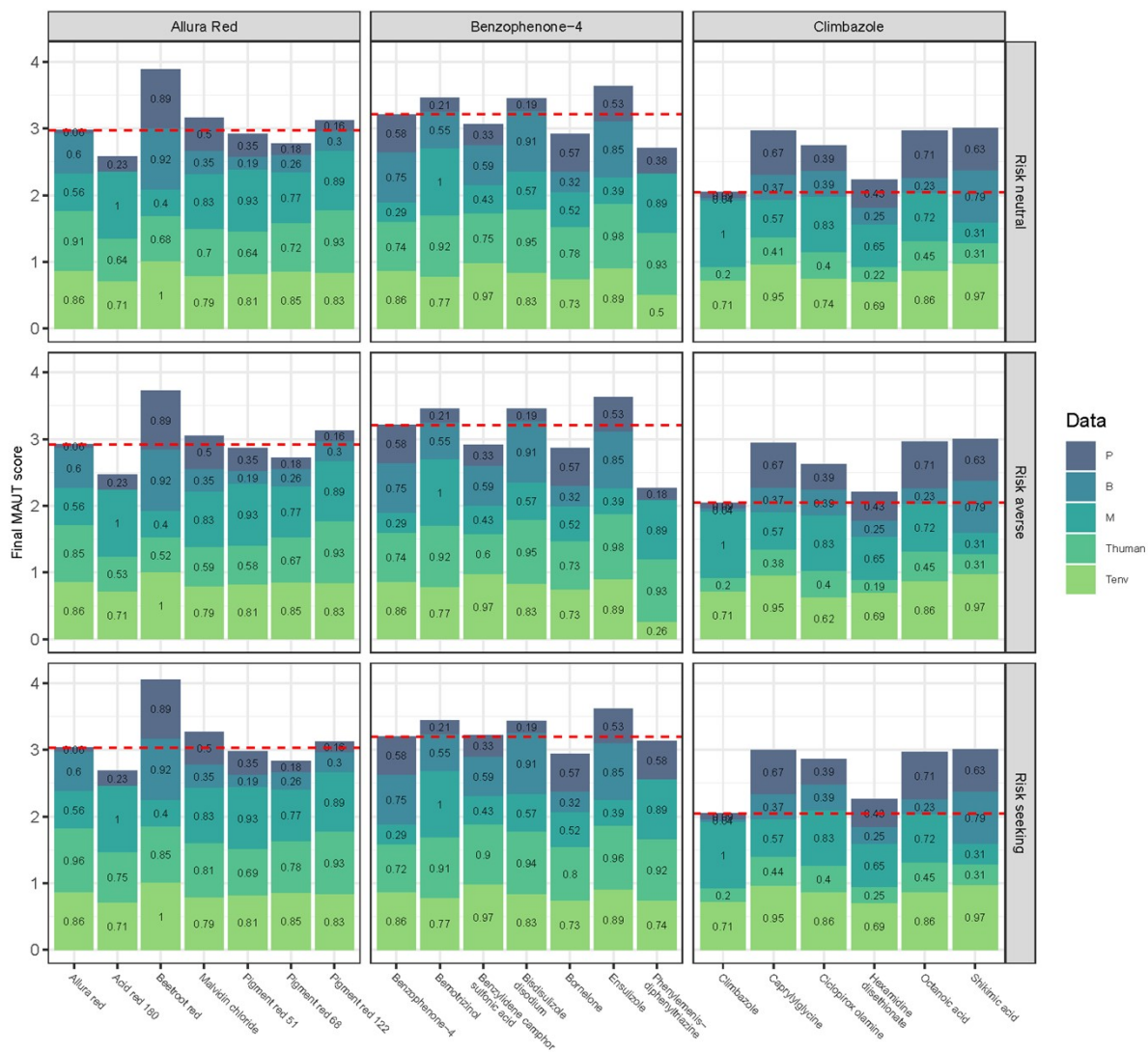
Figure 2 Results of the MAUT assessment of the three different case-study chemicals and shortlisted alternative substances for three different scenarios to treat data gaps. In the risk neutral, risk averse and risk seeking scenario data gaps were assigned a value of 0.5, 0.2 and 0.8, respectively. The red dashed line represent the final PBMT score of the substance to phase out. Every alternatives with a higher PBMT score were considered as safer.

3.3. ELECTREIII

ELECTRE III is an outranking approach that compares different alternatives with each other (Rowley et al., 2012). Three important thresholds for each hazard endpoint are necessary for this method: the indifference threshold (*a*), the preference threshold (*b*), and the veto threshold (*c*). In short, when comparing to alternatives for one hazard endpoint, if the difference between the two alternatives is lower than *a*, then the weight of the hazard endpoint is set to 0; if the difference is greater than *b*, then the full weight of the hazard endpoint considered is awarded to the superior alternative. At last, if the difference is as large as *c*, the ELECTRE III method eliminates the underperforming alternative from contention.

The method to determine the value of each threshold for every hazard endpoint was inspired by Zheng et al. (2019). The thresholds for a specific hazard endpoint were determined based on the difference between the level 0 and the level 1 in the MAUT approach for the endpoint. The equations to determine each threshold are provided below.

$$Indifference_{Endpoint\ A} = 0.25 * (Endpoint\ A_{Level\ 1\ in\ MAUT} - Endpoint\ A_{Level\ 0\ in\ MAUT})$$

$$Preference_{Endpoint\ A} = 0.5 * (Endpoint\ A_{Level\ 1\ in\ MAUT} - Endpoint\ A_{Level\ 0\ in\ MAUT})$$

$$Veto_{Endpoint\ A} = 0.5 * (Endpoint\ A_{Level\ 1\ in\ MAUT} - Endpoint\ A_{Level\ 0\ in\ MAUT})$$

The alternatives were then ranked based on the results of the comparison on P, B, M, and T. In this study it was assumed that every hazard endpoint is equally important and therefore have the same weight in the decision-making. The alternative with a higher rank was considered as safer.

The data used for the MAUT approach are presented in tables S5.3A, 3B and 3C.

Table S5.3A: Data used in the ELECTRE III approach to identify safer alternatives to Allura Red

Table S5.3B: Data used in the ELECTRE III approach to identify safer alternatives to Benzophenone-4

Table S5.3C: Data used in the ELECTRE III approach to identify safer alternatives to Climbazole

Each table are separated into three Excel sheets to present data for the risk neutral, risk averse and risk seeking scenarios.

References

EU, 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) (OJ L 396, 30.12.2006, pp. 1-520). https://doi.org/10.1039/9781782622222-00468

Rowley, H.V., Peters, G.M., Lundie, S., Moore, S.J., 2012. Aggregating sustainability indicators: Beyond the weighted sum. Journal of Environmental Management 111, 24–33. https://doi.org/10.1016/j.jenvman.2012.05.004

US EPA, 2012. Estimation Programs Interface Suite™ for Microsoft® Windows (EPISUITE), v 4.11.