Value and limitation of structure-based profilers to characterize developmental and reproductive toxicity potential

Serena Manganelli[1] · Benoît Schilter[1] · Gabriele Scholz[1] · Emilio Benfenati[2] · Elena Lo Piparo[1]

Abstract

The uncertainty regarding the safety of chemicals leaching from food packaging triggers attention. In silico models provide solutions for screening of these chemicals, since many are toxicologically uncharacterized. For hazard assessment, information on developmental and reproductive toxicity (DART) is needed. The possibility to apply in silico toxicology to identify and quantify DART alerts was investigated. Open-source models and profilers were applied to 195 packaging chemicals and analogues. An approach based on DART and estrogen receptor (ER) binding profilers and molecular docking was able to identify all except for one chemical with documented DART properties. Twenty percent of the chemicals in the database known to be negative in experimental studies were classified as positive. The scheme was then applied to 121 untested chemicals. Alerts were identified for sixteen of them, five being packaging substances, the others structural analogues. Read-across was then developed to translate alerts into quantitative toxicological values. They can be used to calculate margins of exposure (MoE), the size of which reflects safety concern. The application of this approach appears valuable for hazard characterization of toxicologically untested packaging migrants. It is an alternative to the use of default uncertainty factor (UF) applied to animal chronic toxicity value to handle absence of DART data in hazard characterization.

Keywords Developmental and reproductive toxicity · Endocrine activity · Food contact chemicals · In silico · Read-across · Structural alert

Introduction

The uncertainty regarding the safety of food contact materials (FCM) has triggered increasing public, scientific and regulatory attention. The most consistent data gap deals with migrating substances, which originate from impurities in starting raw materials and from reaction and degradation products formed during material manufacture (the so-called Non-intentionally Added Substances, NIAS). For many of

✉ Serena Manganelli
Serena.Manganelli@rd.nestle.com

1 Chemical Food Safety Group, Nestlé Research, Lausanne, Switzerland

2 Laboratory of Environmental Chemistry and Toxicology, Department of Environmental Health Sciences, Istituto di Ricerche Farmacologiche "Mario Negri"-IRCCS, Milan, Italy

these chemicals, no toxicological data is available, and therefore, to assess their potential health risk is highly challenging. Schilter et al. (2014) developed an approach to establish the level of safety concern of chemicals lacking toxicity data. This approach, allowing to prioritize toxicologically untested chemicals, requires quantitative prediction of toxicological endpoints as well as exposure estimate to calculate a margin of exposure (MoE). Up to now it has only relied on prediction of chronic toxicity (such as from 2-year rodent studies), which in general constitutes the most sensitive endpoint (Kroes et al. 2000). However, for a complete hazard assessment, information on developmental and reproductive toxicity (DART) is also needed (such as obtained from two-generation and teratogenicity studies).

DART includes a number of adverse outcomes, such as abnormalities in fetal growth, fetal death, structural dysmorphogenesis, reproductive impairments and cognitive dysfunctions. Importantly, the occurrence of maternal toxicity increases the complexity of the interpretation of such endpoints, since it may indirectly induce developmental effects (Marzo et al. 2016). Experimental research has suggested

that DART may be elicited by a diversity of mechanisms (Wu et al. 2013), although for most chemicals known to produce such types of effects, exact mechanisms of action are not fully elucidated (National Research Council 2000). Activation or/and inhibition of hormone nuclear receptors are well described mechanisms of action of certain endocrine disruptors and are considered as highly plausible molecular initiating events involved in chemical-induced DART (National Research Council 2000; UNEP/WHO 2013). Indeed, hormonal systems play major roles in normal developmental processes and interfering with these during critical windows of development is known to often result in permanent adverse effects later in life (UNEP/WHO 2013). Therefore, endocrine activity such as through nuclear receptor interaction can be considered as a strong alert for DART.

In absence of appropriate DART data, the use of uncertainty factors (UF) (3 or 10) to be applied to chronic animal no observed adverse effect level (NOAEL) has been recommended by governmental bodies, i.e., the US Environmental Protection Agency (EPA). In other framework, the size of such factor has been revised based on the level of incompleteness of available toxicological databases, e.g., when chronic NOAEL is available only in one species (Gadagbui et al. 2005). More recently, Blackburn et al. (2015) proposed that the results of in silico models [i.e., the DART decision tree developed by Wu et al. (2013)] could be valuable in informing the appropriate magnitude of uncertainty regarding developmental and reproductive toxicity.

The present work investigates the possibility to apply in silico toxicology to identify and quantify alert for DART with the ultimate aim to avoid the systematic application of conservative default UF to handle absence of experimental DART data in hazard characterization.

Publicly available predictive in silico tools exist, e.g., models implemented in the VEGA platform (www.vegahub.eu) and in the OECD QSAR Toolbox profilers (www.qsartoolbox.org/). Some of them address developmental toxicity only, others combine developmental and reproductive toxicity, while several refer to endocrine potential (i.e., estrogenic and androgenic potential) (Roncaglioni et al. 2008; Cassano et al. 2010; Marzo et al. 2016; Porta et al. 2016; Manganelli et al. 2019). However, they often only give a qualitative information related to the presence or absence of DART alert. This approach is suitable for hazard identification (identification of intrinsic properties of a chemical to cause harm) but it does not provide any information about hazard characterization (how much is needed to trigger a toxic effect) and even less on its potential health risk. Moreover, some models have a quite restricted domain of applicability in terms of chemical diversity, while others are applicable to a broader chemical space. Due to this large variety of data and chemical diversity their integration and application is very challenging. Aside from (quantitative) structure

activity relationship ((Q)SAR) models, molecular docking is an efficient means for identifying potentially endocrine active chemicals (Vedani et al. 2006; Zhang et al. 2016). Compared to (Q)SAR, this method is more time-consuming.

In this context, the present work was aimed at developing and applying a fast in silico strategy to predict potential developmental and reproductive toxicity of FCM chemicals for which experimental data is missing. It follows basic principles similar to those reported by Blackburn et al. (2015). A stepwise approach, incorporating structural alerts (SA) and docking on nuclear receptors followed by read-across to quantify the alert (i.e., defining a NOAEL/LOAEL), was designed. Before being applied to a large list of FCM chemicals, the value of this new approach was tested in a pilot study to a restricted set of toxicologically uncharacterized FCM substances. Importantly, it assigns a quantitative information (NOAEL/LOAEL) for comparison with exposure to obtain a MoE (ratio between toxic level and exposure). The size of the MoE is a reliable indicator of the level of safety concern. Introducing this DART prediction approach to the decision tree proposed by Schilter et al. (2014) will significantly improve the establishment of safety concern of toxicologically uncharacterized food chemicals and could replace the use of default factor to account for the uncertainties related to the absence of experimental DART data.

Materials and methods

Data set curation

The list of 195 chemicals used to develop and validate this strategy includes plastic and ink FCM migrating substances (Price and Chaudhry 2014). We extended the list of 183 chemicals used in our previous work on mutagenicity prediction (Manganelli et al. 2018) by adding twelve FCM structural analogues identified by Price and Chaudhry (2014). In the curation process, experimental values, reported by Price and Chaudhry (2014), were revised and some new ones were introduced following literature and database searches, e.g., from freely available databases compiled for building (Q)SAR models:

- Experimental information from the database of VEGA Estrogen receptor Relative Binding Affinity (ERRBA) model (Roncaglioni et al. 2008), containing human Estrogen Receptor alpha (ERα) binding, from the Ministry of Economy Trade and Industry (METI) of Japan.
- ER experimental binding activities from the database of the VEGA CERAPP model. Data were curated by the United States Environmental Protection Agency (US EPA) under the Collaborative Estrogen Receptor (ER) Activity Prediction Project (CERAPP) (Judson et al.

2015; Mansouri et al. 2016) derived from 18 in vitro ER high-throughput screening (HTS) assays developed by the US ToxCast and Tox21 programs.

- AR experimental binding activities from the database of CoMPARA models (Manganelli et al. 2019). Structures of these chemicals were curated by US EPA in the framework of the Collaborative Modeling Project for Androgen Receptor (AR) Activity (CoMPARA) derived from 11 AR in vitro HTS assays developed by the US ToxCast and Tox21 programs. In CERAPP and CoMPARA, a chemical was considered as a binder if it was either an active agonist or antagonist. We considered classification data only. For more information about the ToxCast assays and how they were combined see Kleinstreuer et al. (2017).
- The data set collected to build VEGA CAESAR v.2.1.7 model for developmental toxicity (Cassano et al. 2010), combining subsets of information from the Teratogen Information System (TERIS) and US Food and Drug Administration (FDA) guidelines. Both sources contain assessments of human and animal data on potentially teratogenic chemicals.
- The P&G data set with DART effects mediated by nuclear receptors, and by other mechanisms obtained from a detailed literature review (Wu et al. 2013). It also includes effects, which occur in the presence of maternal toxicity. This data set has been used to build the developmental/reproductive models (P&G) implemented in the VEGA platform (Marzo et al. 2016).

The final data set of 195 chemicals included 74 chemicals with experimental data: 14 substances gave at least one evidence for DART and/or endocrine activity, the 60 remaining ones were negative for one or more endpoints.

Curation of chemical structures was performed as described previously (Manganelli et al. 2018). The 74 toxicologically characterized chemicals were used to define our approach, which was then used to evaluate and prioritize the 121 remaining chemicals.

Profilers selected for DART hazard identification

The stepwise approach was developed on the 74 experimentally characterized chemicals with the aim to build a workflow easily and quickly applicable to prioritize long lists of chemicals. Therefore, predictions using several models, most of them freely available in the VEGA platform [CAESAR for developmental toxicity, IRFMN/CERAPP, ERRBA and in the OECD QSAR ToolBox (P&G DART and ER binding profilers)], mentioned in the data set curation section, were evaluated. From a toxicological perspective, to have a consistent coverage of potential DART all these models need to be used. However, since

experimental data were mostly collected from the training sets of these models and due to the relatively small number of chemicals of interest, the selection of only two OECD QSAR Toolbox profilers was sufficient to establish the proof of concept for this pilot study. Indeed, Toolbox profilers are not intended to be used as models per se, due to recognised limitations (e.g., lack of an applicability domain). We selected the following profilers, keeping in mind these weaknesses.

P&G DART v.1.0 scheme Toolbox v 4.1

This profiler of the QSAR Toolbox is an adaptation of the decision tree for identifying chemicals with structural features associated with DART potential, developed by Wu et al. (2013), based on a detailed review of 716 chemicals. These chemicals were grouped into different categories and sub-categories. After running a chemical through the decision tree, the results indicate if the chemical of interest is associated or not with chemical structures known to have DART, or that it has structural features outside the chemical domain of the decision tree. We used the P&G DART to achieve a broad coverage of the relevant endpoints described by Wu et al. (2013), since it was built based on the evidence of both reproductive and developmental effects, mediated by nuclear receptors, e.g., AR and ER, and/or other mechanisms. P&G developed an internal automated version of the DART decision tree, defining a chemical domain of the tree and considering out of domain a chemical if the chemical scaffold structure does not overlap with any of the structures considered in the development of the tree. One limitation of the QSAR Toolbox version is that it does not assess the match of the underlying chemical scaffold of the target compound compared to structures addressed by the DART decision tree. Consequently, determination of whether a chemical is in or outside of the chemical domain of the tree needs to be done manually (Blackburn et al. (2015)). In our study we highlighted these limitations.

Estrogen receptor (ER) binding profiler Toolbox v 4.1

The Toolbox profiling scheme for ER binding classifies chemicals based on their molecular weight (MW) and on structural features, as very strong, strong, moderate and non-binders based on experimental data and literature evidences (Schultz et al. 2002; Gallegos Saliner et al. 2003; Hamblen et al. 2003). This profiler was selected to have a coverage of available experimental data on ER mediated effects, not necessarily linked to DART effects, or for which DART potential has not been yet identified.

VirtualToxLab (VTL, version 5.8)

VTL is an in silico tool for predicting the endocrine activity potential of chemicals. The technology is based on an automated protocol that simulates and quantifies the binding of small molecules towards a series of sixteen proteins, mainly nuclear receptors. The approach is not training set dependent, and therefore, the applicability domain concept cannot be applied with the only restriction that the ligand should have MW less than 500. The approach has been extensively validated (Vedani et al. 2012), and in our study, the predictions were used to confirm ER and AR structural alerts highlighted by the other models. VTL is based on the three-dimensional interactions between the ligand and nuclear receptors, but it does not provide information on if the binding affinity will trigger or not an endocrine effect. Predicted binding affinities were converted in -log units and expressed in the following ranges: i. strong binders: binding affinity ≥ 9 (nanoM); ii. moderate binders; $9 >$ binding affinity ≥ 6 (microM); iii. weak binders: $6 >$ binding affinity ≥ 3 (milliM).

Read-across for DART hazard characterization

To assign a quantitative developmental/reproductive toxic value to chemicals for which a DART alert was identified, read-across was performed. Developmental/reproductive (N) LOAEL values of the most similar chemical belonging to the relevant P&G (sub)category were used for that purpose. Similarity was based on Tanimoto coefficient implemented in KNIME (Berthold al. 2008), ranging from 0 (i.e., maximum dissimilarity) to 1 (i.e., maximum similarity). Similar chemicals were promoted for read-across based on expert judgment.

Reproductive or developmental NOAELs and LOAELs were retrieved from several sources such as: Agency for Toxic Substances and Disease Registry (ATSDR, https://www.atsdr.cdc.gov/toxprofiledocs/index.html), OECD Screening Information DataSet (SIDS, https://www.inchem.org/pages/sids.html, TOXNET (https://toxnet.nlm.nih.gov/newtoxnet/hsdb.htm), EFSA Scientific Reports or Scientific publications (https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2006.76r), European Chemicals Agency (ECHA) Registered substances data (https://echa.europa.eu/web/guest/information-on-chemicals/registered-substances). Preference was given to studies with oral rather than any other exposure routes.

Results

The stepwise approach (shown in Fig. 1) was developed on the 74 experimentally characterized chemicals to first identify and then characterize their DART potential. P&G

DART and ER binding profilers were used for hazard identification followed by docking to confirm nuclear receptor binding mediated mechanisms. A read-across step was then introduced and applied to chemicals identified as positive for their DART potential.

Hazard identification

Structural profiler application on experimentally positive chemicals

The P&G DART scheme correctly categorized seven out of fourteen experimentally positive chemicals, the ER binding profiler classified five of the remaining ones as potential binders. Their combination gave alerts for twelve toxicologically positive substances (see Table 1 and Supplementary Table 1S). Overall, six of the experimentally positive chemicals were associated with ER binding. These substances were identified only by the ER binding profiler (see Table 1), except for BPA, which was associated with nuclear receptor mediated DART according to the P&G DART scheme. VTL confirmed the binding profiles of these chemicals, giving more specific details on binding affinity, also considering different receptor subtypes. For example, unlike the ER binding profiler giving the same label of "very strong binder" for bisphenol F isomers, VTL was able to discriminate between them, highlighting higher affinity (moderate) of 4, 4′-bisphenol F towards ER$_\beta$ and weak binding for 2,2′- and 2,4′- isomers.

Two chemicals were not covered by any of the profilers: 2,6-ditert-butyl-4-(3,5-ditert-butyl-4-hydroxyphenyl)phenol,
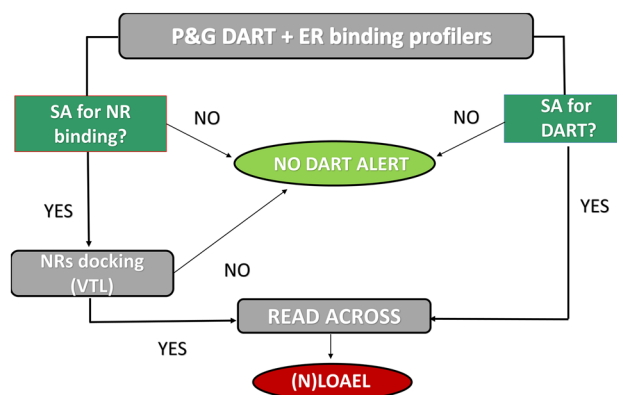

Fig. 1 Stepwise approach for DART hazard identification and characterization. The stepwise approach consists of: (1) Running P&G DART and ER binding Toolbox profilers to identify DART hazard. (2) Chemicals assigned to nuclear receptor binding mediated effects are screened by VTL to confirm binding to Nuclear Receptors (NR) through docking method. (3) For chemicals associated with a SA from P&G scheme and/or ER binding profiler, read-across is performed to assign an experimental quantitative (N)LOAEL value based on the most similar compound contained in the SA class

Table 1 Results of P&G scheme, ER binding profiler and VTL used for hazard identification on experimentally positive characterized chemicals

Chemical associated to structural alerts		Experimental activity				Prediction				
CAS	Name	DART	AR binding activity	ER binding activity	ER binding	P&G DART scheme	ER binding profiler	VTL AR binding	VTL ER$_\alpha$ binding	VTL ER$_\beta$ binding
1137-42-4	4-hydroxybenzophenone	N/A	N/A	N/A	Yes	No	Yes	Moderate	Weak	Moderate
79-14-1	Glycolic acid	Yes	No	No	N/A	Yes	No	N/A	N/A	N/A
104-76-7	2-ethylhexan-1-ol	Yes	No	No	N/A	Yes	No	N/A	N/A	N/A
111-46-6	Diethylene glycol	Yes	No	No	N/A	Yes	No	N/A	N/A	N/A
149-57-5	2-ethylhexanoic acid	Yes	No	No	N/A	Yes	No	N/A	N/A	N/A
2467-02-9	2,2′-Bisphenol F	N/A	Yes	Yes	N/A	No	Yes	Moderate	Weak	Weak
2467-03-0	2,4′-Bisphenol F	N/A	N/A	Yes	N/A	No	Yes	Moderate	Weak	Weak
620-92-8	4,4′-Bisphenol F	N/A	N/A	Yes	Yes	No	Yes	Moderate	Weak	Moderate
3274-28-0	2-propylhexanoic acid	Yes	N/A	N/A	N/A	Yes	No	N/A	N/A	N/A
80-05-7	Bisphenol A	Yes	Yes	Yes	N/A	Yes	Yes	Moderate	Moderate	Moderate
510-15-6	Chlorobenzilate	Yes	Yes	No	N/A	No	No	N/A	N/A	N/A
121-00-6	3-tert-butyl-4-hydroxyanisole	Yes	No	No	No	No	Yes	Weak	Weak	Weak
131-58-8	2-methylbenzophenone	Yes?[a]	Yes?	Yes?	Yes?	Yes	No	N/A	N/A	N/A
128-38-1	2,6-ditert-butyl-4-(3,5-ditert-butyl-4-hydroxyphenyl)phenol	Yes?	Yes?	Yes?	Yes?	No	No	N/A	N/A	N/A

N/A not available (experimental activity) and not applied (VTL)

[a]Highly uncertain/not confirmed/effect not specified

which was labelled as positive by Price and Chaudhry (2014), and chlorobenzilate with a positive AR activity label from CoMPARA. Further literature and database searches did not provide any experimental data for 2,6-ditert-butyl-4-(3,5-ditert-butyl-4-hydroxyphenyl)phenol, and we concluded that the experimental value reported by Price is highly uncertain. Regarding chlorobenzilate, it was used as a pesticide and it is reported as FCM structural analogue by Price, thus it is not specifically a food contact chemical. This chemical was evaluated as a probable human carcinogen by US EPA. Injury to the sperm and atrophy of the testes have been observed in rats exposed to chlorobenzilate in their diet (US EPA 1999), which confirms that this chemical is a false negative.

Structural profiler application on experimentally negative chemicals

We also examined the 60 chemicals which were experimentally negative for DART/estrogenic/androgenic potential. Twelve of them were classified as positive by at least one profiler (nine by P&G DART and three by ER binding profiler) (see Tables 2 and 2S). We examined experimental data of these to understand if their positive labels indicate true misclassification.

Triphenyl phosphate and triethyl phosphonoacetate were categorized by P&G DART profiler as organophosphorus compounds. Triphenyl phosphate, a printing ink compound (Bradley et al. 2013), showed no evidence of toxicity to fertility in an old one-generation study in rat (Welsh et al. 1987) and in a recently performed prenatal developmental toxicity study in rabbit according to ECHA registration data (NOAEL = 200 mg/kg BW/day) (ECHA 2019). It was also negative for in vitro AR binding, according to CoMPARA data. For Triethyl phosphonoacetate, an additive with restricted use in PET (EFSA 2008), there were no direct effects on visceral and/or skeletal development of foetuses in prenatal developmental toxicity study and the repeated-dose toxicity study did not show any effect on fertility, as reported in ECHA registration data (ECHA 2019). Even if in this case the two chemicals falling in this category were not developmental toxicants according to our experimental data search, this is still an important information given by the profiler. Indeed, organophosphates represent a large subcategory of chemicals potentially generating neurodevelopmental toxicity primarily by the inhibition of acetylcholinesterase (AChE) enzyme inhibition (Wu et al. 2013).

4-Methylbenzophenone, a photoinitiator for printing inks in the packaging field (EFSA 2009a), was classified as negative by Price and Chaudhry (2014) and falls in P&G DART category "Toluene and small alkyl toluene derivatives (8a)". Chemicals associated to this category are primarily developmental toxicants, and reprotoxic in the presence of a nitro-group (Wu et al. 2013). Unfortunately, literature searches fail to identify any experimental data which could inform on potential DART activity. Therefore, it was not possible to confirm or refuse the hypothesis that it is a false positive.

Propane-1,2-diol and butane-1,3-diol fall in the P&G DART category "Di-substituted hydrocarbons (24a)". This category is mostly based on evidence of toxicity exerted by chemicals, which are not reproductive or developmental toxicants per se; however, they are converted into active metabolites (Wu et al. 2013). This class does not refer to AR and ER mediated mechanisms, inline with available AR and ER experimental data reported in our database. For propane-1,2-diol, a food additive in the EU, no adverse effects were observed in the available oral reproductive and developmental toxicity studies (EFSA 2018). Similarly, 1,3-butanediol (1,3-butylene glycol), included in the EU register of flavorings substances used in or on foodstuffs, was considered to be non-developmentally toxic (EFSA 2011).

n-Butyl acrylate is classified in the P&G category "Vinyl amide, aldehyde and ester derivatives (21a)", which is based on teratogenic chemicals acting through non-NR mediated DART mechanism, inline with available absence of AR and ER experimental activity. For this chemical, adverse effects were observed only in the presence of maternal toxicity after exceeding the upper limit dose recommended by the respective OECD guidelines according to ECHA registration data. Since the P&G scheme considers also DART effects occurring in the presence of maternal toxicity (Wu et al. 2013), the information available indicates that this chemical may not be misclassified according to the intrinsic P&G DART criteria. However, classifying a chemical as developmental toxicant when the effect is observed at maternally toxic dose is a matter of debate. *n*-Butyl methacrylate and ethyl acrylate fall in the same P&G category as butyl acrylate and similar reasoning may be applied.

The three remaining chemicals were not classified for DART effects by P&G but they are labelled as binders by the ER binding profilers and by VTL. 2-hydroxybenzophenone is a false positive, since it is experimentally negative for ER binding. We did not find further experimental information on DART/endocrine activity for this chemical through databases and literature search.

Octabenzone and 2-aminoanthraquinone are structural analogues of certain FCM chemicals. Octabenzone is experimentally negative for AR and ER binding activity but is classified as binder by ER Toolbox profiler, confirmed by VTL. It cannot be considered a false positive, since experimental negative data refer to binding related to activity, while the ER-binding profiler refers solely to binding. 2-Aminoanthraquinone, profiled as strong ER binder by Toolbox, is classified as weak binder towards ER receptors and as non-binder towards AR receptor by VTL, inline with available

Table 2 Experimentally negative chemicals (for ER/AR developmental or reproductive effects) profiled by P&G scheme, ER binding profilers and VTL

Chemical associated to structural alerts		Experimental activity				Prediction				
CAS	Name	DART	AR binding activity	ER binding activity	ER binding	P&G DART scheme	ER binding profiler	VTL AR binding	VTL ERα binding	VTL ERβ binding
117-99-7	2-Hydroxybenzophenone	N/A	No	N/A	No	No	Yes	Moderate	Weak	Weak
115-86-6	triphenyl phosphate	No	No	N/A	N/A	NC	No	N/A	N/A	N/A
57-55-6	Propane-1,2-diol	No	No	No	N/A	Yes	No	N/A	N/A	N/A
107-88-0	Butane-1,3-diol	No	No	N/A	N/A	Yes	No	N/A	N/A	N/A
867-13-0	Triethyl phosphonoacetate	No	N/A	N/A	N/A	NC	No	N/A	N/A	N/A
141-32-2	n-butyl acrylate	No	No	No	N/A	Yes	No	N/A	N/A	N/A
97-88-1	Butyl methacrylate	No	N/A	N/A	N/A	Yes	No	N/A	N/A	N/A
1843-05-6	Octabenzone	N/A	No	No	N/A	No	Yes	Moderate	Moderate	Weak
117-79-3	2-aminoanthraquinone	N/A	No	N/A	N/A	No	Yes	No	Weak	Weak
134-84-9	4-methylbenzophenone	No?[a]	No?	No?	NO?	Yes	No	N/A	N/A	N/A
80-62-6	Methyl methacrylate	No	N/A	N/A	N/A	Yes	No	N/A	N/A	N/A
140-88-5	Ethyl acrylate	No	N/A	N/A	N/A	Yes	No	N/A	N/A	N/A

NC not covered by the decision tree (but identified)

[a] Highly uncertain/not confirmed/effect or endpoint not specified

in vitro experimental activity. Since in vitro experimental activity of 2-aminoanthraquinone refers to AR and prediction to ER binding, the chemical cannot be considered a false positive.

Hazard characterization

Read-across

For each chemical associated with a P&G DART category and/or with nuclear receptor binding confirmed by VTL, the most similar compound of the relevant P&G category was identified and proposed for read-across. Experimental developmental/reproductive (N)LOAEL data of the most similar chemicals were used to fill the toxicological data gaps of these target chemicals when appropriate. Organophosphates were not considered for read-across, since a specific threshold of toxicological concern is available and can be readily applied (Kroes et al. 2004).

Some chemicals [i.e., bisphenol A (BPA), Glycolic acid, 2-ethylhexan-1-ol, 2-ethylhexanoic acid, 2-propylhexanoic acid] are part of the categories/data collected by Wu et al. (2013), thus the most similar compound is the compound itself (similarity index = 1). Among the remaining chemicals, 2- and 4-methylbenzophenone, falling in the P&G (sub) category, "Toluene and small alkyl toluene derivatives (8a)", have a similarity index of 0.42 and 0.48 with o-xylene and 4-tert-butyltoluene, their most similar chemicals in this P&G category (see Supplementary Tables 1S and 2S). They share only the toluene ring with this chemical and not the methylbenzophenone, i.e., meaning similarity is insufficient to support read-across and may not lead to a fair comparison in read-across. Similarly, octabenzone, 3-tert-butyl-4-hydroxyanisole (3-BHA) and 2-aminoanthraquinone, profiled by ER Binding Profiler share only some functional groups when associated with their most similar chemicals in the P&G category 2b (related to ER and AR binding), which are tamoxifen, BPA and isoproteron, respectively, with similarity values of 0.42, 0.48, and 0.32, respectively. Based on this evidence, for chemicals falling in these two categories, i.e., Toluene and small alkyl toluene derivatives (8a) and ER and AR binding (2b), we established a cut-off similarity value of 0.5 as initial criterion to consider read-across reliable. For most chemicals associated to ER binding confirmed by VTL (i.e., Bisphenol F isomers, 2- and 4-hydroxybenzophenone, 3-BHA) the most similar chemical found in the P&G category 2b (related to ER and AR binding) was BPA (see Tables 1S and 2S). Their similarity values in comparison with BPA were greater than 0.5 except for 3-BHA, which was slightly below this threshold. 2-aminoanthraquinone and octabenzanone are structural analogues of FCM (as reported by Price) and not food contact chemicals as such. For analogues exceeding the expert-based 0.5 cut-off value in these

two classes, we searched for their developmental/reproductive (N)LOAEL values to be assigned to the target chemicals through read-across (see Table 3). In some cases, inhalation values were found, which may be converted into oral values (Schilter et al. 2014). Actual (N)LOAEL values for developmental/reproductive toxicity of target chemicals, where available, are reported to test the approach. However, in most cases experimental designs were different from each other and the highest doses tested were too far from each other to enable a fair comparison between (N)LOAEL values.

Design of a stepwise approach

Based on the results described above, a stepwise approach for DART hazard identification and characterization was designed, as depicted in Fig. 1. Toolbox profilers are the first steps to identify DART hazard. Based on the limited number of chemicals of this study, sequential or parallel application of these two profilers did not affect final classification and results. Chemicals assigned to nuclear receptor binding mediated effects are screened using VTL to confirm binding to nuclear receptors docking method. Considering the limitations of the application of structural profilers, positive ER binding from Toolbox is considered only if confirmed by VTL. For each chemical associated with a SA from P&G scheme and/or ER binding profiler, similarity search to identify the most similar chemical contained in the relevant P&G category is conducted. For SA not identified by P&G DART scheme but identified by the ER binding profiler and confirmed by VTL, similarity search is performed with the compounds contained in the relevant P&G category related to ER and AR binding (2b) (Wu et al. 2013). Read-across is used to assign an experimental quantitative (N)LOAEL value based on the most similar compound contained in the SA class.

Application of the stepwise approach to toxicologically uncharacterized FCM chemicals

Finally, the stepwise approach was applied to the 121 untested chemicals (for DART) starting with profilers' application: overall, alerts were identified for 18 chemicals (see Table 4). The P&G DART scheme raised alerts for 13 out of 121 substances and ER binding profiler for seven of them, two were classified positive by both.

Chemicals predicted as DART positive

Benzanthracene-5,6-oxide, an FCM structural analogue identified by Price, fell in the P&G subcategory "AhR binders. Polycyclic aromatic hydrocarbons (PAHs) (3b-3)". We applied VTL, which confirmed AhR binding as stronger than the other screened receptors. For all chemicals falling

Table 3 Experimentally characterized chemicals promoted for read-across and their most similar chemicals in the relevant P&G category
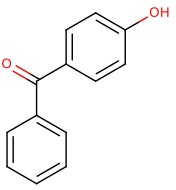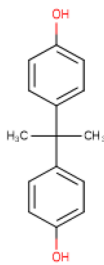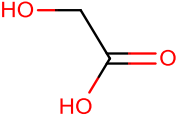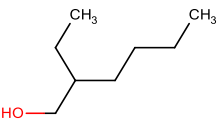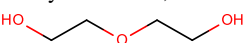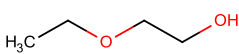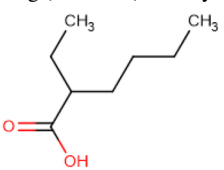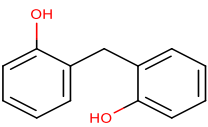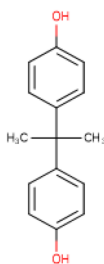
Chemical associated to structural alerts	The most similar in the relevant P&G category

4-benzoylphenol (4-hydroxybenzophenone; *p*-hydroxybenzophenone)**,** CAS 1137–42-4
No DART data

Bisphenol A (BPA), CAS 80-05-7; SI = 0.74
Oral NOAEL (multigeneration studies in rodents) = 5 mg/kgBW/day (EFSA 2015)

2-hydroxyacetic acid (Glycolic acid), CAS 79-14-1

Compound itself (SI = 1)
Oral developmental and maternal toxicity NO(A)EL in rat = 150 mg/kgBW/day (Munley et al. 1999)

2-ethylhexan-1-ol, CAS 104-76-7

Compound itself (SI = 1)
Maternal and fetal oral NOAEL in rat = 130 mg/kgBW/day (McGinty et al. 2010)

2-(2-hydroxyethoxy)ethanol (Diethylene glycol), CAS 111-46-6

Oral developmental and maternal toxicity NO(A)EL in rats (gavage) = 1118 (Ballantyne and Snellings 2005)

Ethylene glycol monoethyl ether (2-etoxyethanol), CAS 110–80-5; SI = 0.94
Oral Reproductive LOAEL in rat (gavage) = 100 mg/kgBW/d (Adedara and Farombi 2010)

Compound itself (SI = 1)
Rat oral NOAEL = 100 mg/kgBW/day (EFSA 2009b)

2-ethylhexanoic acid, CAS 149-57-5

2,2′-Bisphenol F, CAS 2467-02-9
No DART studies

BPA; SI = 0.62
Oral NOAEL (multigeneration studies in rodents) = 5 mg/kgBW/day (EFSA 2015)

Table 3 (continued)
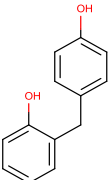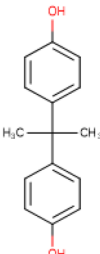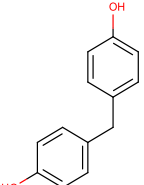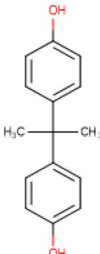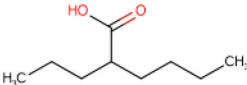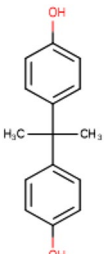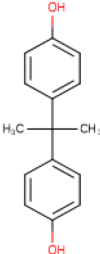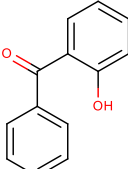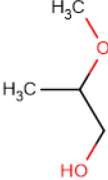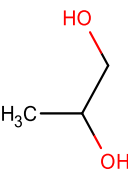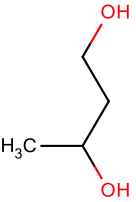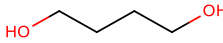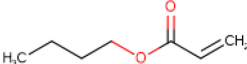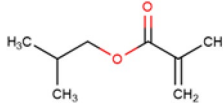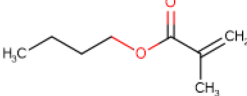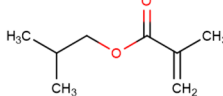
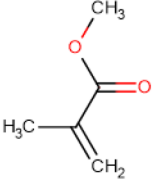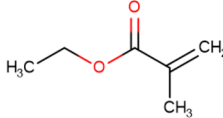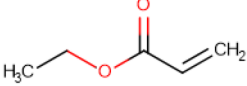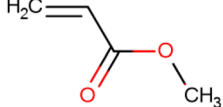| Chemical associated to structural alerts	The most similar in the relevant P&G category
2,4′-Bisphenol F, CAS 2467-03-0 No DART studies	BPA; SI = 0.66
4,4′-Bisphenol F, CAS 620–92-8 No DART studies	BPA; SI = 0.80
2-propylhexanoic acid, CAS 3274-28-0	Compound itself; SI = 1 Oral Developmental LOAEL (rat) = 900 mg/kgBW/day (Narotsky et al. 1994)
Bisphenol A, CAS 80-05-7	Compound itself; SI = 1 Oral NOAEL (multigeneration studies in rodents) = 5 mg/kgBW/day (EFSA 2015)
2-benzoylphenol (2-hydroxybenzophenone), CAS 117-99-7; No DART data	BPA; SI = 0.59; Oral NOAEL (multigeneration studies in rodents) = 5 mg/kg BW/day (EFSA 2015)
Propane-1,2-diol (propylene glycol), CAS 57–55-6 Oral prenatal developmental/reproductive and maternal NOAEL in rat = 1000 mg/kg BW/day (only dose tested) (EFSA 2018)	2-Methoxy-1-propanol, CAS 1589-47-5; SI = 0.64 Inhalation prenatal toxicity LOAEL in rabbit = 225 ppm (Hellwig et al. 1994)

Table 3 (continued)

Chemical associated to structural alerts	The most similar in the relevant P&G category
Butane-1,3-diol (1,3-butylene glycol), CAS 107–88-0	1,4-Butanediol; CAS 110–63-4; SI = 0.83
Oral developmental and maternal NOAEL in rat (gavage) = 706 mg/kg bw (EFSA 2012)	Oral reproductive NOAEL in combined repeat dose and reproductive/developmental OECD TG 422 study in rat = 800 mg/kgBW/d; developmental NOAEL in mice = 600 mg/kg BW/day (OECD 2000)
n-butyl acrylate, CAS 141-32-2	Isobutyl methacrylate, CAS 97–86-9; SI = 0.696
Oral rat reproductive NOAEL in extended one-generation study according to OECD 443 = 150 mg/kg BW/day (ECHA 2019)	Only outdated experimental studies or read-across results
n-butyl methacrylate, CAS 97–88-1	Isobutyl methacrylate, CAS 97–86-9 SI = 0.82
Rat oral developmental/reproductive and maternal NOAEL in combined repeat dose and reproductive/developmental OECD TG 422 study = 300 mg/kg/day (OECD 2009)	
Methyl methacrylate, CAS 80-62-6 Rat oral dev/repro NOAEL = 400 mg/kg BW/day (ANSES 2018)	Ethyl methacrylate, CAS 97-63-2; SI = 0.77 Rat inhalation developmental NOAEC = 600 ppm, LOAEC 1200 ppm (Saillenfait et al. 1999a)
Ethyl acrylate, CAS 140-88-5	Methyl acrylate, CAS 96–33-3 SI = 0.79
Rat inhalation developmental NOAEC = 100 ppm and developmental and maternal LOAEC = 200 ppm (Saillenfait et al. 1999b)	Rat inhalation developmental/maternal NOAEC = 50 ppm, LOAEC = 100 ppm (Saillenfait et al. 1999b)

in "Toluene and small alkyl toluene derivatives (8a)" (sub) category, the most similar chemicals have very low similarity indices when compared to their most similar chemicals, ranging from 0.25 to a higher one of 0.43 only for 3-methylbenzophenone. Similarly, 9,10-epoxystearic acid and 1,2-cyclohexanedione fall in P&G categories "glycidyl ether derivatives (16c)" and "piperazine-, dioxane-, morpholine-, tetrahydrothiopyran-like derivatives and cyclohexanamine (17c)" respectively. Their most similar chemicals are N-butyl glycidyl ether (similarity index = 0.41), and cyclohexylamine (similarity index = 0.27).

Since for two chemicals (labeled only by ER binding profiler), ER binding by Toolbox was not confirmed by VTL, sixteen chemicals were prioritized. Eleven of them are structural analogues of FCM. The remaining five include: (i) [3-(3-tert-butyl-4-hydroxy-5-methylphenyl) propyl]oxyphosphonous acid, which is part of a cyclic ester with the substance 3,3′,5,5′-tetrakis(tert-butyl)-2,2′-dihydroxybiphenyl, intended to be used as an antioxidant in the manufacture of polyolefins for contact with all types of foodstuffs (EFSA 2009c); (ii) 4-(4-Methylphenylthio)

Table 4 Toxicologically uncharacterized chemicals, P&G DART, ER binding profilers, VTL classification and the most similar chemicals in the relevant P&G category

CAS	Name	P&G DART	ER binding profiler	VirtualToxLab binding confirmation	The most similar in the relevant P&G category
116-83-6	anthraquinone, 1-amino-4-methoxy	Not known[a]	Strong binder	Weak (ERs)	Tamoxifen (SI=0.32)
129-43-1	1-hydroxyanthracene-9,10-dione	Not known	Strong binder	Weak (ERs)	BPA (SI=0.45)
790-60-3	Benzanthracene-5,6-oxide	AhR binders, PAHs (3b-3)	Non binder	Moderate (AhR)	Benzo[a]pyrene (SI=0.4)
10048-13-2	8-hydroxy-6-methoxy-3a,12c-dihydro-7h-furo[3′,2′:4,5]furo[2,3-c]xanthen-7-one	Not known	Strong binder	Moderate (ERs)	Tamoxifen (SI=0.29)
85-83-6	1-({2-methyl-4-[(2-methyl-phenyl)diazenyl]phenyl}diazenyl)naphthalen-2-ol	Toluene and small alkyl derivatives (8a)	Strong binder	Not binding (ERs)	4-tert-Butyltoluene (SI=0.26)
10254-99-6	Anthraquinone, 2-hexanoyl-1,3,6,8-tetrahydroxy	Not known	Very strong binder	Not binding (ERs)	N/A
137-17-7	2,4,5-trimethylaniline	Not known	Weak binder	Not binding (ERs)	N/A
13394-86-0	2′,3-dimethyl-(1,1′-biphenyl)-4-amine	Toluene and small alkyl derivatives (8a)	Moderate binder	Weak (ERs)	o-xylene (SI=0.30)
2426-07-5	1,2,7,8,diepoxyoctane	Glycidyl ether derivatives (16c)	Non binder	N/A	N-Butyl glycidyl ether (SI=0.36)
2443-39-2	9,10-epoxystearic acid	Glycidyl ether derivatives (16c)	Non binder	N/A	N-Butyl glycidyl ether (SI=0.41)
765-87-7	1,2-cyclohexanedione	piperazine-, dioxane-, morpholine-, tetrahydrothiopyran-like derivatives and cyclohexanamine (17c)	Non binder	N/A	cyclohexylamine (SI=0.27)
643-65-2	3-methylbenzophenone	Toluene and small alkyl derivatives (8a)	Non binder	N/A	m-xylene (SI=0.43)
–	4-{4-[(4-methylphenyl)sulfanyl]phenyl}benzal-dehyde	Toluene and small alkyl derivatives (8a)	Non binder	N/A	p-xylene (SI=0.27)
80-39-7	*N*-Ethyl-*p*-toluenesulfon-amide	Toluene and small alkyl derivatives (8a)	Non binder	N/A	o-xylene (SI=0.27)
83846-85-9	4-(4-Methylphenylthio) benzophenone	Toluene and small alkyl derivatives (8a)	Non binder	N/A	4-tert-Butyltoluene (SI=0.31)
1077-56-1	*n*-ethyl-2-methylbenzene-sulfonamide	Toluene and small alkyl derivatives (8a)	Non binder	N/A	o-xylene (SI=0.27)
82–28-0	1-amino-2-methylanthra-cene-9,10-dione	Toluene and small alkyl derivatives (8a)	Non binder	N/A	4-tert-Butyltoluene (SI=0.25)
–	[3-(3-tert-butyl-4-hydroxy-5-methylphenyl)propyl] oxyphosphonous acid	Toluene and small alkyl derivatives (8a)	Non binder	N/A	4-tert-Butyltoluene (SI=0.29)

SI similarity index

[a]Not known precedent DART potential

benzophenone, (iii) *N*-ethyl-*p*-toluenesulfonamide and (iv) 3-methylbenzophenone, which are printing ink compounds (Bradley et al. 2013); (v) 4-{4-[(4-methylphenyl) sulfanyl]phenyl}benzaldehyde, for which no information on application or source was found. For these chemicals, similarity was too low to apply read-across only within P&G categories (see Supplementary Table 3S).

Discussion

The approach proposed is an improved strategy for translating DART alerts into a quantitative value to be used to establish a level of safety concern according to risk assessment principles. It represents a proof of concept study based on a limited number of FCM substances with the ultimate aim to be applied later to a significantly larger number of compounds. Here we applied profilers related to DART covering different mechanisms of action represented by the SA embedded in the P&G scheme and by NR-binding prediction. The selection of models used was tested on a limited number of chemicals. It can be refined at any time if required, i.e., through inclusion of new models for hazard identification. The stepwise approach can be considered as a workflow that can be applied using other (even proprietary) validated DART models both for SA/QSAR as well docking. In addition, read-across will be more reliably applicable if based on a bigger library of either food contact chemicals and/or other chemicals with appropriate toxicity data available.

The Toolbox profilers are not intended to be used as independent in silico models. However, for documented DART chemicals the combination of the P&G DART scheme and ER binding profilers followed by VTL delivered almost a full coverage of positive hits on the small set of molecules we considered, leading to only one misclassification for the FCM analogue chlorobenzilate. For this chemical, AR binding activity was not identified by the Toolbox, highlighting the need to apply at least an additional model to cover specifically AR binding activity. Consequently, the inclusion of more in silico models covering a broader range of NRs in the first screening phase is advisable. This will be addressed in the next stage, when this stepwise approach will be applied to a larger number of substances. On the other hand, for experimentally negative chemicals, positive toxicity profiling does not indicate misclassification by default. Indeed, AR or ER binding may not always be associated with developmental and/or reproductive toxicity.

The first screening phase of chemicals allowed to reduce the number of chemicals to be tested with docking, which is a more sophisticated but more time-consuming technique requiring sharper competences. For some toxicologically characterized and uncharacterized chemicals docking highlighted stronger binding affinities on specific receptors. This may help to define the best analogues to conduct read-across. Additionally, docking did not confirm ER binding for two uncharacterized structural analogues of FCMs and, therefore, constituted the driver for actual classification.

Most importantly, we propose a read-across approach allowing to translate structural alerts into a quantitative (N)LOAEL to be compared with exposure estimate to obtain a margin of exposure (MoE). The size of the MoE informs on the level of safety concern. The read-across approach is based on structural similarity and is, furthermore, backed up by mechanistic plausibility. Chemical similarity search was performed among chemicals belonging to the same P&G category of the target mostly sharing the same mode of action (MoA). Importantly, similarity analysis highlighted some weaknesses of the P&G DART scheme profiler implemented in the QSAR Toolbox. In particular, chemicals labeled as "Toluene and small alkyl derivatives (8a)" have low similarity indices when compared to their analogues and, most importantly, contain more than 50% of other significant structural moieties in addition to the toluene structural scaffold.

In a number of cases, prioritized chemicals were in the database of the model, thus their quantitative reproductive/developmental (N)LOAEL were directly used to quantify structural alerts. For most of the others, read-across could be applied using the most suitable analogue of the proper class. For the remaining, read-across assessment by applying a one-to-one comparison between target and the most similar chemical in the relevant P&G category was not feasible due to a lack of appropriate experimental data for the comparator. It is important to note that read-across can be performed using a category approach, extending the comparison to relevant chemicals (with suitable DART data) falling in the same category but not contained in the P&G database. Finally, when read-across is not applicable on chemicals presenting clear DART alerts, the possibility to apply an UF of 10 on chronic L(N)OAEL is still possible, as proposed by Blackburn et al. (2015). These authors showed that chemicals identified by the original DART tree (Wu et al. 2013) as being related to structures with known DART toxicity would potentially have lower DART NOAELs compared to their respective repeated-dose toxicity NOAELs than structures that lacked this association. They proposed to apply an UF of 10 to chronic NOAELs for chemicals giving exact match with structures with DART precedent according to P&G decision tree, or an UF of three for chemicals not matching with DART alert but in domain of the decision tree. In our case the most conservative UF of 10 would be recommended, since the P&G DART profiler implemented in Toolbox does not provide info on applicability domain.

The results described in the present paper indicate that methodologies (e.g., Schilter et al. 2014) aimed at establishing level of safety concern of chemicals without toxicological data may significantly benefit from integrating the approach proposed in the present paper. Indeed, bringing evaluation of DART potential would allow a more complete hazard assessment complying better with risk assessment principles. However, limitations and uncertainties have to be kept in mind. The models/profilers limitations are highlighted in the material and method section. One remaining uncertainty is the adequacy of the selected models/profilers

to cover all DART endpoints/mechanisms. The P&G scheme has been documented to cover an array of alerts and mechanisms of DART. In our investigation, the addition of tools to predict interactions with estrogen receptor was considered sufficient to identify correctly all chemicals of our dataset documented to produce DART effects, probably due to the relatively low number of active substances in our dataset.

The uncertainties linked to the application of the proposed approach are more difficult to characterize. Schilter et al. (2014) observed and concluded that establishing safety concern based on an integration of high quality and relevant in silico toxicology predictions was unlikely to be significantly more uncertain than that based on experimental data (Schilter et al. 2014). Such conclusion may apply to the present approach dealing with DART hazard characterization, although uncertainties associated with read-across may be higher. The main uncertainties associated with read-across concern the number and degree of similarity of the chemical analogues, as well as the uncertainties related to the toxicological data found for the analogues.

Recent trends on read-across extended the concept of similarity not only to the chemical structure, but also to other features, related to the toxicological profile and toxicokinetics aspects, including metabolism (Schultz et al. 2015; Patlewicz et al. 2017). The use of SA in read-across supports this concept. Thus, read-across can be extended in this direction. Any improvement in read-across and in the set of in silico models will have an impact in two directions: (1) to increase elements for the evaluation, covering potential gaps; (2) allowing a refinement of the uncertainty of the overall strategy. With a reference to the EFSA Guidance on weight-of-evidence (EFSA 2017), the inclusion of these models will represent additional lines of evidence. According to this guidance, the different lines of evidence must be evaluated considering their relevance, reliability and consistency. In the case of read-across, if similar substances are not so similar to a given target, as in several cases we discussed, their relevance is very low, and they can be disregarded. We discussed also about the reliability of the experimental values for the case of the read-across. Another element to consider is the consistency between multiple lines of evidence. The acceptability of the level of uncertainty is not an absolute value. The integration of results of multiple in silico models and read-across within the conceptual scheme of the EFSA guidance (2017) has been discussed by Benfenati et al. (2019).

Conclusions

The present approach offers the possibility to identify and assess chemicals for which hazards are driven by developmental and/or reproductive toxicity. For such chemicals,

MoE based on DART could be used for safety assessment. The major drawback of the present approach is the scarcity of developmental/reproductive data which may reduce the possibility to conduct read-across. However, in cases of chemicals with DART alerts but not eligible for read-across because of the absence of appropriate analogues, chronic toxicity values may still be applied but with the application of a factor to take into account the uncertainty regarding DART effects.

Compliance with ethical standards

References

Adedara IA, Farombi EO (2010) Induction of oxidative damage in the testes and spermatozoa and hematotoxicity in rats exposed to multiple doses of ethylene glycol monoethyl ether. Hum Exp Toxicol 29:801–812. https://doi.org/10.1177/0960327109360115

ANSES (2018) Substance evaluation conclusion as required by REACH Article 48 and evaluation report for Methyl methacrylate EC No 201–297–1 CAS No 80–62–6. https://echa.europa.eu/documents/10162/c92faa6c-7134-fc58-5266-5b373cdc9286. Accessed 11 Dec 2019

Ballantyne B, Snellings WM (2005) Developmental toxicity study with diethylene glycol dosed by gavage to CD rats and CD-1 mice. Food Chem Toxicol 43:1637–1646. https://doi.org/10.1016/j.fct.2005.05.005

Benfenati E, Chaudhry Q, Gini G, Dorne JL (2019) Integrating in silico models and read-across methods for predicting toxicity of chemicals: a step-wise strategy. Environ Int 131:105060. https://doi.org/10.1016/j.envint.2019.105060

Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Thiel K, Wiswedel B (2008) KNIME: the konstanz information miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R (eds) Data analysis, machine learning and applications. Studies in classification, data analysis, and knowledge organization. Springer, Berlin, pp 319–326

Blackburn K, Daston G, Fisher J, Lester C, Naciff JM, Rufer ES, Stuard SB, Woeller K (2015) A strategy for safety assessment

of chemicals with data gaps for developmental and/or reproductive toxicity. Regul Toxicol Pharmacol 72:202–215. https://doi.org/10.1016/j.yrtph.2015.04.006

Bradley EL, Stratton JS, Leak J, Lister L, Castle L (2013) Printing ink compounds in foods: UK survey results. Food Additiv Contam: Part B 6:73–83. https://doi.org/10.1080/19393210.2012.725774

Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D, Benfenati E (2010) CAESAR models for developmental toxicity. Chem Cent J 4:S4. https://doi.org/10.1186/1752-153X-4-S1-S4

CIR (2018) Safety assessment of triphenyl phosphate as used in cosmetics. https://www.cir-safety.org/sites/default/files/tripho06202018FR.pdf. Accessed 4 Dec 2019

ECHA (2019) https://echa.europa.eu/. Accessed 4 Dec 2019

EFSA (2008) Scientific Opinion of the Panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on substances in food contact materials. EFSA J 6:699

EFSA (2009) EFSA statement on the presence of 4-methylbenzophenone found in breakfast cereals. EFSA J 7:243. https://doi.org/10.2903/j.efsa.2009.243r

EFSA (2009) Scientific opinion of the panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the European Commission on Flavouring Group Evaluation 4: 2-Ethylhexyl derivatives from chemical group 2. EFSA J 7:929. https://doi.org/10.2903/j.efsa.2009.929

EFSA (2009) Scientific Opinion of the Panel on food contact materials, enzymes, flavourings and processing aids (CEF) on 23rd list of substances for food contact materials. EFSA J. https://doi.org/10.2903/j.efsa.2009.1028

EFSA (2011) Scientific Opinion on the evaluation of the substances currently on the list in the Annex to Commission Directive 96/3/EC as acceptable previous cargoes for edible fats and oils—Part I of III. EFSA J 9:2482. https://doi.org/10.2903/j.efsa.2011.2482

EFSA (2012) Scientific Opinion on Flavouring Group Evaluation 10, Revision 3 (FGE.10Rev3): Aliphatic primary and secondary saturated and unsaturated alcohols, aldehydes, acetals, carboxylic acids and esters containing an additional oxygenated functional group and lactones from chemical groups 9, 13 and 30. EFSA J 10:2563. https://doi.org/10.2903/j.efsa.2012.2563

EFSA (2015) Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs: Executive summary. EFSA J 13:3978. https://doi.org/10.2903/j.efsa.2015.3978

EFSA (2017) Scientific Opinion on the guidance on the use of the weight of evidence approach in scientific assessments. EFSA J 15:4971. https://doi.org/10.2903/j.efsa.2017.4971

EFSA (2018) Re-evaluation of propane-1, 2-diol (E 1520) as a food additive. EFSA J 16:5235. https://doi.org/10.2903/j.efsa.2018.5235

Gadagbui B, Zhao J, Maier A, Dourson M (2005) The scientific rationale for deriving database and toxicodynamic uncertainty factors for reproductive or developmental toxicants. Toxicology Exellence for Risk Assessment, Cincinnati, USA

Gallegos Saliner A, Amat L, Carbó-Dorca R, Schultz TW, Cronin MT (2003) Molecular quantum similarity analysis of estrogenic activity. J Chem Inf Comput Sci 43:1166–1176. https://doi.org/10.1021/ci034014a

Hamblen EL, Cronin MT, Schultz TW (2003) Estrogenicity and acute toxicity of selected anilines using a recombinant yeast assay. Chemosphere 52:1173–1181. https://doi.org/10.1016/S0045-6535(03)00333-3

Hellwig J, Klimisch HJ, Jäckh R (1994) Prenatal toxicity of inhalation exposure to 2-methoxypropanol-1 in rabbits. Fundam Appl Toxicol 23:608–613. https://doi.org/10.1006/faat.1994.1147

Judson RS, Magpantay FM et al (2015) Integrated model of chemical perturbations of a biological pathway using 18 in vitro highthroughput screening assays for the estrogen receptor. Toxicol Sci 148:137–154. https://doi.org/10.1093/toxsci/kfv168

Kleinstreuer NC, Ceger P, Watt ED, Martin M, Houck K, Browne P, Thomas RS, Casey WM, Dix DJ, Allen D et al (2017) Development and validation of a computational model for androgen receptor activity. Chem Res Toxicol 30:946–964. https://doi.org/10.1021/acs.chemrestox.6b00347

Kroes R, Galli C, Munro I, Schilter B, Tran L, Walker R, Würtzen G (2000) Threshold of toxicological concern for chemical substances present in the diet: a practical tool for assessing the need for toxicity testing. Food Chem Toxicol 38:255–312. https://doi.org/10.1016/S0278691599001209

Kroes R, Renwick AG, Cheeseman M et al (2004) Structure-based thresholds of Toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Food Chem toxicol 42:65–83. https://doi.org/10.1016/j.fct.2003.08.006

Manganelli S, Schilter B, Benfenati E, Manganaro A, Piparo LE (2018) Integrated strategy for mutagenicity prediction applied to food contact chemicals. ALTEX 35: 169–178. 10.14573/altex.1707171.

Manganelli S, Roncaglioni A, Mansouri K, Judson RS, Benfenati E, Manganaro A, Ruiz P (2019) Development, validation and integration of in silico models to identify androgen active chemicals. Chemosphere 220:204–215. https://doi.org/10.1016/j.chemosphere.2018.12.131

Mansouri K, Abdelaziz A et al (2016) CERAPP: collaborative estrogen receptor activity prediction project. Environ Health Persp 124:1023–1033. https://doi.org/10.1289/ehp.1510267

Marzo M, Kulkarni S, Manganaro A, Roncaglioni A, Wu S, Barton-Maclaren TS, Lester C, Benfenati E (2016) Integrating in silico models to enhance predictivity for developmental toxicity. Toxicology 370:127–137. https://doi.org/10.1016/j.tox.2016.09.015

McGinty D, Scognamiglio J, Letizia CS, Api AM (2010) Fragrance material review on 2-ethyl-1-hexanol. Food Chem Toxicol 48:S115–S129. https://doi.org/10.1016/j.fct.2010.05.042

Munley SM, Kennedy GL, Hurtt ME (1999) Developmental toxicity study of glycolic acid in rats. Drug Chem Toxicol 22:569–582. https://doi.org/10.3109/01480549908993168

Narotsky MG, Francis EZ, Kavlock RJ (1994) Developmental toxicity and structure-activity relationships of aliphatic acids, including dose-response assessment of valproic acid in mice and rats. Fundam Appl Toxicol 22:251–265

National Research Council (2000) Scientific frontiers in developmental toxicology and risk assessment. National Academies Press, Washington DC

OECD (2000) Screening Information Dataset (SIDS) Initial Assessment Report. 1,4-butanediol. CAS No 110–63–4. SIAM 10 (Tokyo, March 15–17, 2000). https://hpvchemicals.oecd.org/ui/handler.axd?id=e38c211c-0860-4f74-bd71-8c6da255604d Accessed 11 Dec 2019

OECD (2009) Screening information data set for short chain alkyl methacrylates. https://hpvchemicals.oecd.org/ui/SIDS_Details.aspx?id=319e0a7e-feac-4468-824d-f9661b37a8ac. Accessed 11 Dec 2019

Patlewicz G, Helman G, Pradeep P, Shah I (2017) Navigating through the minefield of read-across tools: a review of in silico tools for grouping. Comput Toxicol 3:1–18. https://doi.org/10.1016/j.comtox.2017.05.003

Porta N, Roncaglioni A, Marzo M, Benfenati E (2016) QSAR Methods to Screen Endocrine Disruptors. Nucl Receptor Res. 10.11131/2016/101203.

Price N, Chaudhry Q (2014) Application of in silico modelling to estimate toxicity of migrating substances from food packaging. Food Chem Toxicol 71:136–141. https://doi.org/10.1016/j.fct.2014.05.022

Roncaglioni A, Piclin N, Pintore M, Benfenati E (2008) Binary classification models for endocrine disrupter effects mediated

through the estrogen receptor. SAR QSAR in Environ Res 19:697–733. https://doi.org/10.1080/10629360802550606

Saillenfait AM, Bonnet P, Gallissot F, Peltier A, Fabries JF (1999a) Developmental toxicities of methacrylic acid, ethyl methacrylate, *n*-butyl methacrylate, and allyl methacrylate in rats following inhalation exposure. Toxicol Sci 50:136–145. https://doi.org/10.1093/toxsci/50.1.136

Saillenfait A, Bonnet P, Gallissot F, Protois JC, Peltier A, Fabries JF (1999b) Relative developmental toxicities of acrylates in rats following inhalation exposure. Toxicol Sci 48:240–254. https://doi.org/10.1093/toxsci/48.2.240

Schilter B, Benigni R, Boobis A, Chiodini A, Cockburn A, Cronin MTD, Lo Piparo E, Modi S, Thiel A, Worth A (2014) Establishing the level of safety concern for chemicals in food without the need for toxicity testing. Regul Toxicol Pharm 68:275–296. https://doi.org/10.1016/j.yrtph.2013.08.018

Schultz TW, Sinks GD, Cronin MT (2002) Structure-activity relationships for gene activation oestrogenicity: evaluation of a diverse set of aromatic chemicals. Environ Toxicol 17:14–23. https://doi.org/10.1002/tox.10027

Schultz TW, Amcoff P et al (2015) A strategy for structuring and reporting a read-across prediction of toxicity. Regul Toxicol Pharmacol 72:586–601. https://doi.org/10.1016/j.yrtph.2015.05.016

UNEP, WHO (2013) State of the science of endocrine disrupting chemicals-2012. World Health Organization, Geneva

US EPA (1999) Integrated risk information system (IRIS) on chlorobenzilate. National Center for Environmental Assessment, Office of Research and Development, Washington, DC. https://www.epa.gov/sites/production/files/2016-09/documents/chlorobenzilate.pdf. Accessed 1 Aug 2019

Vedani A, Dobler M, Lill MA (2006) The challenge of predicting drug toxicity in silico. Basic Clin Pharmacol Toxicol 99:195–208. https://doi.org/10.1111/j.1742-7843.2006.pto_471.x

Vedani A, Dobler M, Smieško M (2012) VirtualToxLab—a platform for estimating the toxic potential of drugs, chemicals and natural products. Toxicol Appl Pharmacol 261:142–153. https://doi.org/10.1016/j.taap.2012.03.018

Welsh JJ, Collins TFX, Whitby KE, Black TN, Arnold A (1987) Teratogenic potential of triphenyl phosphate in Sprague–Dawley (Spartan) rats. Toxicol Ind Health 3:357–369

Wu S, Fisher J, Naciff J, Laufersweiler M, Lester C, Daston G, Blackburn K (2013) Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem Res Toxicol 26:1840–1861. https://doi.org/10.1021/tx400226u

Zhang J, Begum A, Brännström K, Grundström C, Iakovleva I, Olofsson A, Sauer-Eriksson AE, Andersson PL (2016) Structure-based virtual screening protocol for in silico identification of potential thyroid disrupting chemicals targeting transthyretin. Environ Sci Technol 50:11984–11993. https://doi.org/10.1021/acs.est.6b02771