



HAL
open science

The way forward for assessing the human health safety of cosmetics in the EU: Proceedings of Workshop 2

Vera Rogiers, Emilio Benfenati, Ulrike Bernauer, Laurent Bodin, Patience Browne, Nicolas Cabaton, Qasim Chaudhry, Pieter Jan Coenraads, Emanuela Corsini, Matthew Dent, et al.

► **To cite this version:**

Vera Rogiers, Emilio Benfenati, Ulrike Bernauer, Laurent Bodin, Patience Browne, et al.. The way forward for assessing the human health safety of cosmetics in the EU: Proceedings of Workshop 2. Archives of Toxicology, In press, <10.1007/s00204-025-04275-x>. <hal-05447922>

HAL Id: hal-05447922

<https://hal.science/hal-05447922v1>

Submitted on 12 Feb 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



The way forward for assessing the human health safety of cosmetics in the EU: Proceedings of Workshop 2

SCCS - Scientific Committee on Consumer Safety¹ · Vera Rogiers² · Emilio Benfenati³ · Ulrike Bernauer⁴ · Laurent Bodin⁵ · Patience Browne⁶ · Nicolas Cabaton⁷ · Qasim Chaudhry⁸ · Pieter Jan Coenraads⁹ · Emanuela Corsini¹⁰ · Matthew Dent¹¹ · Corie Ellison¹² · Janine Ezendam¹³ · Eric Gaffet¹⁴ · Corrado Lodovico Galli¹⁵ · Effrosyni Katsanou¹⁶ · Aglaia Koutsodimou¹⁷ · Henriqueta Louro¹⁸ · Abdulkarim Najjar¹⁹ · Alicia Paini²⁰ · Eirini Panteri²¹ · Christophe Rousselle²² · Nikaeta Sadekar²³ · Andreas Schepky²⁴ · Maciej Stępnik²⁵ · Wolfgang Uter²⁶ · Tamara Vanhaecke²⁷ · Natalie von Goetz²⁷ · Susan Wijnhoven²⁸

Received: 27 November 2025 / Accepted: 9 December 2025

© The Author(s) 2025

Abstract

Under the European Cosmetic Regulation, safety assessments of cosmetics and their ingredients must be conducted without the use of animals. This regulatory requirement poses a number of challenges, as validated alternative methods are only available for some of the toxicological endpoints that are typically considered in standard human health risk assessments. Despite significant progress since the ban in 2013, particularly in the development of New Approach Methodologies (NAMs) for local and acute toxicity, and for mutagenicity/genotoxicity, there remains an urgent need for non-animal test methods to assess systemic toxicity, which often becomes evident after repeated or long-term exposure. Currently, no validated animal-free alternatives are available for assessing sub-acute, sub-chronic and chronic toxicity, carcinogenicity, developmental/reproductive toxicity, or for a major part of toxicokinetics. In response to these challenges, the Methodology Working Group of the Scientific Committee on Consumer Safety organised a dedicated workshop in December 2024 to discuss advances in the application of Next Generation Risk Assessment (NGRA) as a strategic animal-free approach for the safety assessment of cosmetic ingredients. The workshop focused on a number of important key issues for the practical application of NAMs and NGRA, their regulatory acceptance and identification of possible (partial) solutions to overcome existing limitations.

Keywords Cosmetics · New approach methodology · NAM · Animal-free method · Replacement method · Alternative method · Next generation risk assessment · NGRA · Safety assessment · Safety evaluation · Scientific Committee on Consumer Safety · SCCS

Abbreviations

ADME Absorption, metabolism, distribution and excretion
AOP Adverse outcome pathways
AUC Area under the curve
BER Bioactivity exposure ratio

C_{max} Maximal blood concentration
CMR Carcinogen, mutagen, reproductive toxic
DART Developmental and reproductive toxicology
EATS Estrogen, androgen, thyroid, steroidogenesis
ER Estrogenic receptor
ECHA European Chemicals Agency

Follow up Report of the Scientific Committee on Consumer Safety (SCCS) Methodology Workshop on Next Generation Risk

Assessment December 10, 2024, Brussels.

Disclaimer This report reflects the content of the workshop and the presentations given by the speakers. It does not necessarily reflect the opinion of the SCCS, EFSA and the European Commission.

Extended author information available on the last page of the article

EFSA	European Food Safety Authority
EMA	European Medicines Agency
FDA	Food and Drug Administration
GD	Guidance document
GIVIMP	Good in vitro method practices
HTPK	High throughput pharmacokinetics
ICCR	International Cooperation on Cosmetics Regulation
ICCS	Advancing animal-free safety assessments for cosmetics
InhalTTC	Inhalation TTC
ITEM	Institute for Toxicology and Experimental Medicine
IVIVE	In vitro to in vivo extrapolation
JRC	Joint Research Centre
LOAEL	Lowest observed adverse effect level
MoA	Mode of action
MoS	Margin of safety
MOIE	Margin of internal exposure
NAM	New approach methodology
NGRA	Next generation risk assessment
NOAEL	No observed adverse effect level
NOEC	No observed effect concentration
NoG	Notes of guidance
OECD	Organisation for Economic Cooperation and Development
PBK	Physiologically-based kinetic
PBPK	Physiologically based pharmacokinetics
PBTK	Physiologically based toxicokinetics
PK	Pharmaco-kinetic
PoD	Point of departure
PoD sys	Point of departure (systemically obtained)
QAF	QSAR assessment framework
QMRF	QSAR model reporting format
QPRF	QSAR prediction reporting format
QRRF	QSAR Results Reporting Format
(Q)SAR	(Quantitative) structure–activity relationship
RAx	Read across
RIFM	Research Institute of Fragrance Materials
SCCS	Scientific Committee on Consumer Safety
TG	Test guideline
(i)TTC	(Internal) thresholds of toxicological concern
WHO	World Health Organization
WoE	Weight of evidence

Introduction

The safety of cosmetics in the EU is ensured by using safe ingredients through a standard regulatory risk assessment that looks at exposure as well as a number of local and systemic toxicological endpoints. In the past, this was done by

using data obtained from experimental animals, but animal testing for the purpose of assessing the safety of cosmetic products and ingredients has not been permitted in the EU since the implementation of the EU Cosmetics Regulation (EU) No 1223/2009 (EU 2009). Animal-free alternative methods, usually termed New Approach Methodologies (NAMs), have to be used instead. Consequently, the Scientific Committee on Consumer Safety (SCCS) has reworked its ‘Notes of Guidance for the testing of cosmetic ingredients and their safety evaluation’ (NoG), 12th Revision, to give priority to animal-free NAMs (SCCS/1647/2022). These consist primarily of validated in vitro methods—such as those for skin and eye irritation, skin sensitisation, phototoxicity, dermal absorption, mutagenicity/genotoxicity, and certain endocrine activity assays—which are often complemented by in silico Quantitative Structure–Activity Relationship {(Q)SAR} models and read-across (RAx). Also, a physicochemical characterisation of the test substances is needed to cover data on purity/impurity profiles, stability, and other key physicochemical parameters such as lipophilicity, pKa, etc. Also, of particular interest are Physiologically Based (Pharmaco)Kinetic {PB(P)K} and Toxicokinetic (PBTK) models to provide data on toxicokinetic aspects (SCCS/1647/2022). Historical in vivo data may still be used, as long as the legislative deadlines for animal testing bans are respected (11 September 2004 for cosmetic products; and 11 March 2009 for all cosmetic ingredients, with the exception of long-term tests done outside Europe, for which 11 March 2013 is the end date). Regulatory decisions on safety/non-safety are based on the Weight of Evidence (WoE) approach, which integrates multiple lines of evidence. This evaluation can be strengthened by additional data from omics-methodologies, high-throughput screening and, where available, human clinical and/or biomonitoring studies. Eventually, other non-animal testing data may be added. For cosmetics ingredients and products, there is generally a regulatory acceptance for validated animal-free methodologies, which is the current situation for local toxicity and acute effects, mutagenicity/genotoxicity, but not for systemic and chronic endpoints for which no validated animal-free NAMs are currently available, apart from some in silico (Q)SAR models and read-across. While the lack of validated animal-free NAMs makes many types of risk assessment more challenging, this is particularly true in the case of nanomaterials, as currently there are no validated NAMs for particle-related aspects.

Next Generation Risk Assessment (NGRA), an alternative risk assessment methodology based on animal-free NAMs, has been introduced in the cosmetic field (Berggren et al. 2017) (Fig 1). Nine principles were established by the International Cooperation on Cosmetics Regulation (ICCR) (Dent et al. 2018, 2021). A number of publications on NGRA

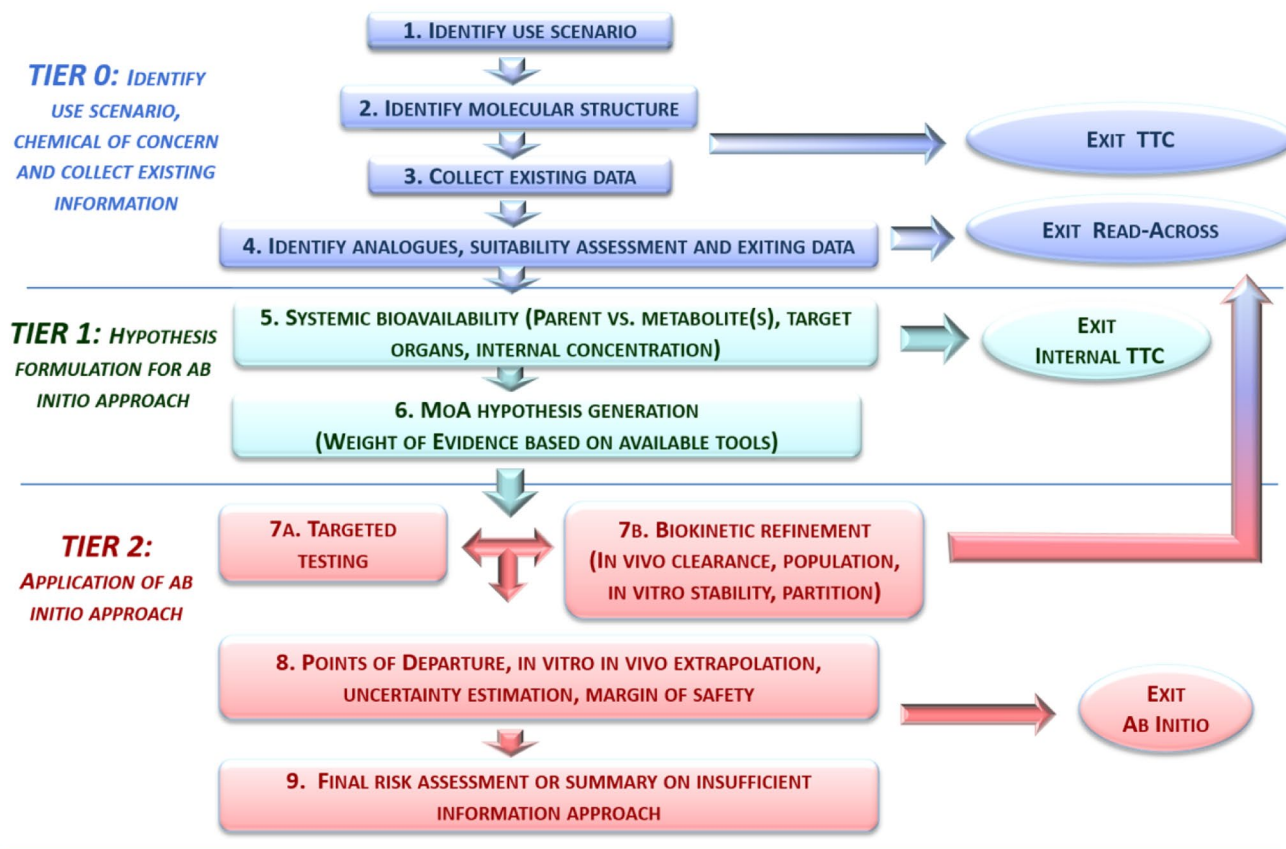


Fig. 1 Framework of New Generation Risk Assessment (NGRA) (taken from Berggren et al. 2017 and Dent et al. 2018). TTC: Threshold of Toxicological Concern; MoA: Mode of Action. Copyright from Elsevier, first published in Computational Toxicology, 4, 2017

case studies have appeared in the open literature, dealing in particular with *ab initio* application of NGRA to assess systemic toxicity of cosmetic ingredients. However, these studies primarily involved data-rich substances (Gilmour et al. 2023; Baltazar et al. 2020; Middleton et al. 2022; Cable et al. 2024; Sepehri et al. 2025), for which SCCS Opinions were available. To further stimulate the acceptability and practical use of NGRA, the SCCS recently created a safe harbour for sharing confidential data/information. Making use of this new possibility, Cosmetics Europe recently submitted two dossiers in parallel to the SCCS, one based on standard risk assessment (SCCS/1660/2023) and the other on an *ab initio* NGRA approach for systemic toxicity.

This article presents the proceedings of the second dedicated Workshop that was organised by the Methodology Working Group of the SCCS in December 2024 to discuss the progress made since the first Workshop in February 2019 in using NGRA as a strategic, animal-free approach for assessing the systemic toxicity of cosmetic ingredients (Rogiers et al. 2020).

Bottlenecks hindering regulatory uptake of NGRA in cosmetic risk assessment: some of these challenges were on the agenda for the second Workshop and up for discussion

- (i) PBK plays an important role in hypothesis generation in an *ab initio* NGRA approach for systemic toxicity. Until now SCCS has not adopted clear criteria on how to reliably verify the outcomes of PBK modeling when no in vivo data is available.
- (ii) In Vitro-In Vivo Extrapolation (IVIVE) was initially developed to estimate Absorption, Distribution, Metabolism, and Excretion (ADME) properties by scaling in vitro experimental results to in vivo contexts (Kulkarni et al. 2021). Also, the approach formed a reliable basis for PBK modeling and enabled a bottom-up approach to toxicokinetic assessment (Kuepfer et al. 2016). As progress has now been made leading to quantitative IVIVE, (Q)IVIVE, it is important to look at what impact this will have on the animal-free risk assessments of cosmetics and their ingredients.
- (iii) Most scientific articles in the open literature concerned with animal-free NGRA of cosmetic ingredients

are based on NAMs that have not yet been validated, which may hinder their consideration in the regulatory decision-making process. OECD Test Guideline (TG) 34 on ‘*Validation and international acceptance of new and updated test methods for hazard assessment*’ is currently under refinement/revision and is also taken up in the ‘Roadmap towards phasing out animal testing for chemical safety assessments’ (Walder et al. 2025). Besides the fact that both processes are complex and time consuming, a significant number of tests still need to be validated once the international criteria become available. This need is taken up in the EU Roadmap towards phasing out animal testing for chemical safety assessments, but a rigorous action plan still needs to be worked out (EC 2025).

- (iv) Another point of interest in animal-free risk assessment is evaluating the actual content of the toxicodynamics toolbox. Which tools are available and how they could be applied in alternative risk assessment such as NGRA for systemic toxicity are among the questions that need to be addressed.
- (v) When NGRA is used for animal-free systemic risk assessment, decisions are often based on the outcome of the bioactivity exposure ratio (BER). A BER greater than 1 suggests that the bioactivity threshold surpasses estimated human exposure, indicating that the substance is potential safe (Health Canada 2021). A BER below 1 suggests that bioactivity may occur at the estimated human exposure, which means that safety cannot be assured unless the exposure is lowered (either through refinement of estimates or by lowering the levels of intended use). The interpretation of BER values is under discussion and needs to be agreed upon for risk assessments carried out by the SCCS. Further evidence is needed to establish firm regulatory thresholds for different types of NAMs (Wambaugh et al. 2018).
- (vi) The Threshold of Toxicological Concern (TTC) and internal TTC (iTTC) are potential exit points in the NGRA frameworks (Fig. 1). The question arises whether TTC can be applied to exposure routes other than dermal, e.g. inhalation exposure. The same question will come up when iTTC data becomes available.
- (vii) Despite efforts made at the EU level to define some critical rules for the application of (Q)SAR models for toxicological purposes, there remains a risk of misinterpretation when applying the QSAR Toolbox, or other (Q)SAR models and RAX, potentially leading to incorrect decisions. Further clarification is also needed in this field.

Various issues, as depicted in the flow chart below, were discussed by the invited speakers, stakeholders, and experts

from the Cosmetics Industry, Academia and regulatory authorities, including the European Commission, at the second Workshop.

Criteria for accepting PBK modeling for new substances in the absence of in vivo data?

PBK modelling has emerged as a fundamental tool in the assessment of the biokinetics of chemicals and in extrapolating in vitro data to in vivo by reverse dosimetry (Bell et al. 2018; Wambaugh et al. 2018). PBK models simulate the ADME processes within the human body, offering valuable insights into how substances behave in vivo (Andersen 1995). The main components of PBK models include compartments representing different tissues (e.g., liver, kidneys, brain) and the processes that occur within and between these compartments. PBK modeling involves the use of mathematical differential equations to describe the behavior of a chemical in the body. These models are based on the physiological characteristics of the human body, including tissue volumes, blood flow rates, and enzyme activity. These models are particularly useful for situations where traditional clinical, human and/or animal data are unavailable or insufficient. One such scenario is the assessment of new chemical substances for which no in vivo data exists, often due to limited human exposure and lack of in vivo animal information (Paini et al. 2021). In these cases, PBK models can be used to predict the ADME processes and biokinetics, and potential toxicity of chemicals based on available in vitro data, chemical structure, and other factors. These types of models may also rely on literature values for physiological parameters and use computational methods to predict the fate of the chemical substance in the body. However, there is a need to establish criteria for PBK model acceptance in the context of new chemicals for which in vivo data is lacking for model development and evaluation.

Proposal for the acceptance of key criteria for PBK models

(i) Data quality and availability

The quality of input data is one of the primary determinants of PBK model reliability. In the absence of in vivo data, other sources of data become crucial, such as in vitro experiments, computational predictions, and literature values. Physicochemical properties, metabolic stability, and partition coefficients should be measured/achieved following standardised guidelines (such as OECD TGs). In the absence of TGs, the OECD Good In Vitro Method Practices

(GIVIMP) should be followed (OECD 2018). GIVIMP provides guidance for test method developers and end users of resulting data on key elements of in vitro methods (OECD 2018). In situations where no direct data is available, computational approaches such as molecular docking, QSAR, or machine-learning models may be used to estimate these parameters (Madden et al. 2019). Human tissue volumes, blood flow rates, and organ-specific metabolic enzyme activities are essential for accurate model predictions. These are usually standardised and as such can be retrieved from literature. If in vivo data are unavailable for a particular chemical, these parameters can be borrowed from similar compounds (e.g. RAX), or general physiological knowledge or by allometric scaling.

In vitro data should be extrapolated to in vivo by applying scaling factors available in the literature (Wetmore et al. 2012, 2013, 2015). While IVIVE can be useful, it has limitations, and accurate scaling requires a thorough understanding of the in vitro models and conditions.

(ii) Model Structure and Complexity

The structure of the PBK model should be appropriate for the chemical substance being studied. For new chemicals, model complexity must be carefully considered. PBK models can vary from simple compartmental models to more sophisticated multi-compartment models (Pletz et al. 2020). In the absence of in vivo data, the following factors must be carefully balanced:

- simplicity versus detail;
- number and type of compartments; for new chemicals, tissue distribution based on chemical properties can guide their choice;
- choice of absorption model based on available in vitro data for solubility, permeability, transporter interactions;
- incorporation of known or predicted metabolic pathways, enzyme kinetics, and elimination routes; predictions could be based on chemical structure and available metabolic data from similar compounds.

(iii) Calibration and validation

Calibration and Validation are essential for ensuring that the PBK model provides reliable predictions. If no in vivo data are available for the new chemical substances, data from structurally similar compounds can be used for comparison. While not a direct validation, this approach helps to check the model's internal consistency and evaluates its predictive power. It is important to stress that as new data becomes available, the PBK model should be refined. For example, the model can be updated with new in vitro or in

silico data, or it may be adjusted based on expert judgment or external knowledge.

One of the key aspects of model validation is assessing the uncertainty and variability in the predictions. Sensitivity analysis allows researchers to identify which parameters most affect the model's output (Najjar et al. 2022). For new chemicals with limited data, sensitivity analysis is particularly important to ensure that predictions are robust and reliable. Parameters with the greatest uncertainty, such as metabolic rate constants or tissue partition coefficients, should be clearly identified. Sensitivity analysis helps to assess the impact of these uncertainties on predictions of exposure, efficacy, and toxicity. Uncertainty quantification involves the use of statistical methods, such as Monte Carlo simulations or Bayesian approaches, to propagate the uncertainty in input parameters through the model. This approach provides a range of possible outcomes rather than a single deterministic result, offering a more realistic prediction of chemical behavior.

Regarding the 'Proof of Predictivity', it should not be solely based on a specific value or threshold. Najjar et al. (2022) referenced a statement from the World Health Organisation (WHO 2010), which noted that '*In PBK modelling, predictions that are, on average, within a factor of 2 of the experimental data have often been considered adequate*'. This threshold is also present in the SCCS NoG (SCCS/1647/2022). In the specific case considered here, this value might be too conservative, since biological factors, at a minimum, account for a factor of 3 (WHO 2005). For more generic PBK models, which require a broader applicability domain, a stringent limit of 2 could become unrealistic. For instance, Punt et al. (2022) found that variations in in vivo C_{max} values can range from 1.3 to fivefold, and even as much as 16-fold. U.S. EPA researchers, analysing in vivo toxicokinetic (TK) data, showed that replicate in vivo measurements are often within a factor of two of each other around 80% of the time. These findings suggest that a more flexible range for the deviation between predicted and observed kinetics could be more appropriate (Cook et al. 2022). It is crucial to acknowledge that both model simulations and experimental data are subject to uncertainty, which must be considered to allow for a fair comparison (WHO 2010). Since establishing a rigid limit for model predictions is challenging, alternative qualitative approaches have been proposed. These include analysing the discrepancies between the model and experimental data and, based on understanding the reasons for these differences, determining whether the model is suitable for extrapolation and under which conditions (Najjar et al. 2022).

(iv) PBK model Platforms

To develop and run PBK model simulations, various programming languages and platforms can be used:

- MATLAB (<https://www.mathworks.com/products/matlab.html>),
- R/RStudio (<https://cran.r-project.org/bin/windows/base/>),
- Python (<https://www.python.org/>),
- C++ (<https://isocpp.org/>),
- Fortran (<https://fortran-lang.org/>),
- Berkely Madonna (<https://berkeley-madonna.myshopify.com/>).
- MCsim (<https://www.gnu.org/software/mcsim/>).

Additionally, user-friendly generic PBK platforms exist, like:

- GastroPlus (<https://www.simulations-plus.com/software/gastroplus/>),
- PKSim (<https://www.open-systems-pharmacology.org/>),
- SimCyp (<https://www.certara.com/software/simcyp-pbpk/>),
- TKplate (<https://r4eu.efsa.europa.eu/>),
- HHTK (<https://cran.r-project.org/web/packages/hhtk/index.html>)

This list of programming languages and platforms for PBK modeling is not exhaustive. The choice of platform largely depends on the modeler's goals, budget, and expertise. When selecting a platform or programming language, key factors to consider include ease of use, flexibility, scalability, availability of libraries, and computational efficiency. The decision should be based on the specific needs of the PBK model development, for instance, the complexity of the model, available resources, and the required performance. Agencies should allow submissions using any coding language or platform, while recommending clear and transparent documentation, easy accessibility of the tool, and maintaining a dialogue with risk assessors. Open access and open source should be the ideal solution. Submitters should provide all relevant evidence and information to ensure the results can be properly assessed and interpreted, considering that at present not all agencies have the means to evaluate PBK models from different platforms. When venturing into the development or application of PBK models, modelers should engage in dialogue with agencies to find common ground for the application of these tools in risk assessment.

(v) Regulatory considerations

Regulatory acceptance of PBK models is crucial for their use in decision-making processes, particularly in the absence of *in vivo* data. Regulatory agencies, such as the Food and Drug Administration (FDA), European Medicines Agency (EMA), and others, require a rigorous evaluation of the PBK model's quality and its ability to predict human biokinetics and toxicity, and to this end, they use qualification systems (FDA 2018; EMA 2018). In addition, Regulatory Agencies require comprehensive documentation of the PBK model, including detailed descriptions of the model structure, assumptions, parameter values, and input data (Tan et al. 2020). Such transparency ensures that the model can be evaluated by external experts. The model should undergo peer review by independent experts to assess its scientific validity and applicability to the specific chemicals being evaluated.

In 2021, the OECD published a Guidance Document, GD 331, describing a scientific workflow for characterising and validating PBK models developed using *in vitro* and *in silico* data (OECD 331 2021). The GD 331 also includes a template for documenting the model characteristics, and a checklist for evaluating the quality of a model. It also highlights the principles, criteria and tools laid out in the OECD PBK model GD, with the aim of facilitating the dialogue between model developers and risk assessors (Paini et al. 2021). These templates have also been taken up in the SCCS Notes of Guidance, 12th Revision (Appendix 13 of SCCS/1647/2022).

Practical application of IVIVE in cosmetic ingredient risk assessment

(i) IVIVE and quantitative use

As mentioned before, IVIVE is a key methodology providing a framework for translating experimental findings from *in vitro* systems into meaningful predictions about biological and toxicological processes in living organisms, including humans. The approach plays a dual role in toxicokinetic and toxicodynamic assessments (Breen et al. 2021), offering an alternative to traditional animal testing for regulatory risk assessments (Najjar et al. 2022). Over the last decade, the number of VIVE-related publications has increased significantly, highlighting its importance in NGRA (Chang et al. 2022b). Recently, (Q)IVIVE was introduced to quantitatively translate *in vitro* bioactivity concentrations into external doses, through reverse dosimetry (Najjar et al. 2022). By integrating PBK models (Louisse et al. 2017), (Q)IVIVE predicts internal concentrations over time in humans or animals, facilitating the identification of Points of Departure

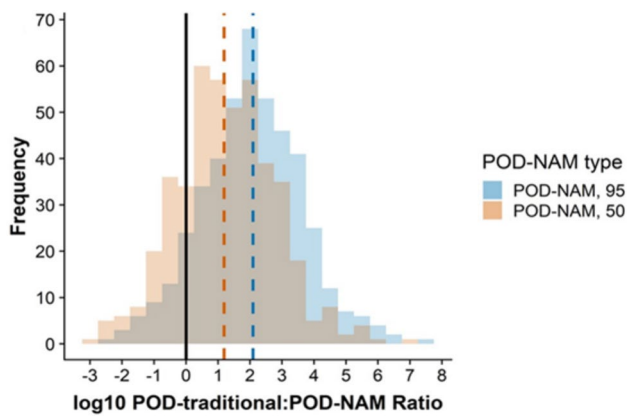


Fig. 2 Log₁₀PoD ratio distribution illustrates for the PoD-NAM, 95 and 50 percentiles. The medians of the Log₁₀PoD ratio distributions are shown as dashed lines for PoD-NAM, 95 and 50 percentiles, figure adapted from Paul-Friedman et al. (2020)

(PoDs) for hazard characterisation or defining safe use concentrations of cosmetic ingredients. On the other hand, PBK modeling can be applied through forward dosimetry to provide information about relevant *in vitro* test concentrations or apply interim iTTC for risk assessment (Najjar et al. 2023). An example by Paul-Friedman et al. (2020) compared PoDs derived from NAMs using a generic high-throughput model with traditional hazard information for 448 chemicals. The study reported the distribution of PoD ratios, showing that NAM-derived PoDs are 10 to 100 times more conservative than traditional *in vivo* PoDs (Fig. 2). Depending on the complexity and the extent of the physiological details, two PBK modeling approaches can be applied. The first is a generic high-throughput modeling such as High Throughput Toxicokinetics (HTTK) and High-Throughput Pharmacokinetics (HTPK) (Naga et al. 2022; Pearce et al. 2017), which is designed for early-stage research, and screening chemicals for potential risks. The second approach involves full PBK models, which are essential for regulatory submissions and detailed risk assessments. Full PBK modeling requires extensive data for parameterisation and is computationally intensive and time-consuming. However, it offers higher accuracy, although comparable accuracy can be achieved with a well-parameterised generic mode (Naga et al. 2022).

(ii) Case Studies

(A) The implementation of (Q)IVIVE in risk assessment of cosmetic ingredients was explored through several cases studies including its role in hazard characterisation (Najjar et al. 2024c, d). Genistein was used as a read-across compound to daidzein, as both interact with estrogen receptor alpha (ER α) and beta (ER β). Legacy *in vivo* data for no observed adverse effect level (NOAEL) was converted using a PBK model to estimate the corresponding internal

exposure (C_{max}). The estimated C_{max} was comparable with the observed *in vitro* PoD derived from CALUX assay on ER α (Najjar et al. 2024c). For daidzein, reverse dosimetry of the no observed effect concentration (NOEC) to its ER α interaction has estimated the corresponding NOAEL at 4.1 mg/kg bw/day, comparable to the reported *in vivo* rats NOAEL at 5 mg/kg bw/day (SCCS/1641/2022). The study highlighted the correlation between the nominal and free test concentrations in the *in vitro* toxicodynamics assays, accounting for non-specific binding and cellular exposure, and the need to refine the observed NAM-based PoDs. Additionally, the extrapolation revealed that the free plasma concentration at the NOAEL dose, rather than the total plasma concentration, correlated with the derived *in vitro* PoD (Najjar et al. 2024c, d). These results support the knowledge that free concentrations represent the active form of a drug or chemical in organisms and raise the need for further research to identify plasma concentrations relevant for risk assessment. In the risk assessment of cosmetic ingredients, (Q)IVIVE can be applied to determining safe exposure levels of compounds, e.g. daidzein and benzyl salicylate (Ebmeyer et al. 2024; Najjar et al. 2024c). In these examples, NAMs-based PoDs derived for the parent ingredients, or their relevant metabolites were extrapolated to estimate safe cosmetic use concentrations. The safety of daidzein in a body lotion was evaluated using the NGRA workflow to define the maximum use concentration, considering the *in vitro* PoD at ER α . A bottom-up human PBK model was developed to estimate the safe external dose for dermal exposure. Since daidzein undergoes phase II metabolism as its major route of elimination, the risk assessment was driven by the parent compound rather than its metabolites. The model incorporated skin first-pass metabolism (Géniès et al. 2025a,b), population variability, and uncertainty of the *in vitro* measurements. The plasma concentration-time profile of daidzein was simulated after repeated use of the body lotion (Najjar et al. 2024d). Model evaluation using legacy data for genistein demonstrated a strong correlation between observed and estimated pharmacokinetic parameters, supporting the model's validity.

(B) Benzyl salicylate, used in face creams at 0.5%, was evaluated for safety using NGRA *ab initio*. Dermal penetration studies revealed intensive skin metabolism, generating several metabolites: salicylic acid, benzoic acid, benzyl alcohol, and hippuric acid (Ebmeyer et al. 2024). Any traces of benzyl salicylate that reach the blood circulation are rapidly metabolised in the liver, producing these metabolites along with additional phase II metabolites. *In vitro* data on the fraction unbound and metabolic stability were generated for benzyl salicylate and its metabolites. A bottom-up human PBK model was developed to simulate the plasma concentration-time profile of benzyl salicylate and the identified

metabolites. The case study identified salicylic acid as the ‘toxdriver’ with the highest internal exposure. Salicylic acid was further tested using toxicodynamics bioassays to obtain the lowest PoDs. A probabilistic PBK model was developed, incorporating probable distributions of the input parameters and population variability. The model estimated the range of internal dose metrics (C_{max} and Area Under the Curve, AUC) following repeated applications of the face cream. The results were used to calculate a so-called “Bioactivity/Exposure Ratio” (BER) based on the PoDs obtained from toxicogenomics assays (Ebmeyer et al. 2024).

(C) An additional example illustrates the implementation of IVIVE/PBK models and advanced modeling methods to provide a better understanding of the toxicokinetics of various salicylate esters (Najjar et al. 2024a). The study estimated their contributions to the aggregated exposure to the common metabolite salicylic acid. It was estimated by integrating a so-called ‘metabolism factor’ into the exposure modeling (Najjar et al. 2024a). To achieve this, a well-stirred model has been used for predicting hepatic clearance from a liver S9 assay, incorporating the fraction of unbound to account for the effect of plasma protein binding and non-specific binding on intrinsic clearance. The observed *in vitro* half-life derived from the liver S9 assay was translated to the *in vivo* half-life in humans using high-throughput PBK modeling. The predicted half-life values were used to calculate the metabolism factor that estimates the converted amount of an ester to salicylic acid within 24h.

(iii) Remaining Issues

Challenges remain in understanding the role of *in vitro* bio-kinetics, variability of parameters and uncertainties, and model evaluation in the absence of *in vivo* data. Currently, robust model evaluation including adherence to good modeling practices with standardised physiological and scaling parameters, efficient sensitivity and uncertainty analyses, and human health protective assumptions to address data gaps are critical for boosting confidence in prediction and fostering regulatory acceptance (Najjar et al. 2022, 2024b).

Readiness criteria: concept, structure and coverage of template proposed by JRC for GD 34 revision

The OECD GD 34 ‘Guidance Document on the validation and international acceptance of new or updated test methods for hazard assessment’ is currently being revised by the OECD, reflecting progress made and addressing outstanding challenges. Against this background, a subgroup was charged with the development of a template for assessing

the readiness of a test method for different steps of the validation process. The template, once finalised and agreed, will be annexed to the revised OECD GD 34 and will provide a comprehensive framework for evaluating the status of development, reliability and relevance of *in vitro/in chemico/ex vivo* testing methods.

The template proposal, developed by the JRC, is based on several existing resources (e.g., EURL ECVAM submission process, Bal-Price et al. 2018; GIVIMP 2018; Krebs et al. 2019; ReadEDTest by Crouzet et al. 2023, Holzer et al. 2023; Petersen et al. 2023; Van der Zalm et al. 2022). The template considers only essential criteria that must be fulfilled before a method enters either validation, transferability or peer-review. It was suggested that the proposed criteria could be also useful for the purposes of the SCCS, which is to judge the validity of non-animal and non-validated methods before they can be used to evaluate risk assessments of cosmetic ingredients.

(i) Template structure and content

Focus lies on individual *in vitro/in chemico/ex vivo* methods. It is endpoint-agnostic and can be applied to methods covering any toxicity endpoint. However, it does not cover *in vivo* and *in silico* methods. The template is structured around four columns:

1. Categories of readiness criteria,
2. Definition of each criterion,
3. Actual criteria in the form of questions to be answered as Yes/No/Not applicable/Unknown,
4. Context to better understand the questions, including references and examples.

The criteria are grouped into five categories:

1. Test method definition and description,
2. Data management, evaluation, and interpretation,
3. Test method reliability,
4. Test method relevance,
5. Data integrity and quality assurance.

Each category includes several sub-criteria. Particular emphasis has been given to the description of the test system, including basic information, quality controls, characterisation of the test system, acceptance criteria, and metabolic competence. Other important elements that need to be addressed include information on the control (and reference) items used in the method, acceptance criteria for the results, and assessment of interferences such as cytotoxicity. In addition, the template includes criteria for assessing various aspects of the relevance of the test method, such

as the biological relevance of the test system (e.g. how the test system is relevant to human toxicity), the mechanistic relevance of the measured endpoints (e.g. why the test endpoint is relevant to an adverse effect), and the regulatory relevance of the test method which is of paramount importance for demonstrating that the biology and mechanisms covered by the test method are indeed relevant to address a regulatory concern.

(ii) Key elements and applications

The template includes essential criteria for evaluating the readiness of a method for validation, transferability, and peer review. It provides a simple ‘Yes/No/Not relevant/I don’t know’ answer format.

Toxicodynamics toolbox in animal-free risk assessment of cosmetic ingredients²⁰

In compliance with the animal testing ban for cosmetics, the identification of the potential mode of action and the corresponding PoDsys has to be evaluated by appropriate animal-free NAMs. The use of a toxicodynamics toolbox containing suitable assays is key. As such, cheminformatics & in silico toxicodynamics tools are part of the toolbox and can be combined with toxicogenomics and a cell stress panel, a pharmacology profiling panel and functional tests, which could also cover developmental and reproductive toxicity (DART) endpoints.

(i) Case Studies

The (Q)IVIVE examples, described under point 4, with focus on systemic toxicity and NAMs were carried out under the umbrella of the Cosmetics Europe Long Range Science Strategy (LRSS). They used a toxicodynamic toolbox showing that a combination of NGRA, RAX and ab initio case studies can be valuable for the determination of the PoDsys of a compound under study.

a. Using a 10-step RAX approach as a case study (Alexander-White et al. 2022), the toxicodynamics toolbox was applied for daidzein to determine the PoDsys based on results of Estrogen, Androgen, Thyroid, Steroidogenesis (EATS) assays combined with a cell stress panel, a pharmacology profiling panel, and toxicogenomics assays. The most relevant endpoint for daidzein was derived from the ER α assay, with a lowest observed effective concentration of 100 ± 0.0 nM, leading to an in vitro PoDsys of 33 nM (Najjar et al. 2024d).

b. In the ab initio benzyl salicylate case study, the toxicity driver was the main metabolite salicylic acid, for which PoDsys was determined. It was significantly directed by toxicogenomics (with a NOEL of 10.6 μ M). In addition, DART was assessed based on in silico models and indications from Toxicity Forecasting (ToxCastTM) using a ReproTrackerTM Assay combined with a cell stress panel and pharmacology profiling panel (Ebmeyer et al. 2024).

c. Application of toolbox components makes it possible to generate a relevant PoDsys. If needed, additional tools or assays can be added, as was the case for the UV-filter benzophenone-4, where the primary human proximal tubule cells model (aProximateTM Assay-Ready Plates for transporter studies) was introduced. By applying NAMs in an NGRA framework, the safety of a substance in a cosmetic product can be properly assessed. This way of working is even more conservative (more protective) than traditional risk assessments.

(ii) Further developments

The pharmacology profiling panel, originally starting with 44 assays, was enlarged to 83 assays (Burbank et al. 2024). As such, this toolbox panel became a great support in capturing toxicodynamic processes. It is of paramount importance to note, however, that toxicogenomics testing must be standardised. In this field, the International Collaboration on Cosmetics Safety (ICCS) has launched promising projects in cooperation with universities, regulatory bodies and industry. The ever-widening possibilities of predicting the toxicological effects of chemical substances by cheminformatics, machine learning and other sophisticated techniques enable precise focusing on critical pathways. These can then be further investigated using existing functional tests, such as DART or EATS assays to yield relevant PoDsys values. A closer look to the ReproTracker[®] assay shows 85% concordance (same for specificity and sensitivity), which is a better performance than earlier used assays (Jamapoor et al. 2022). In this context, also the highly advanced devTOX quickPredictTM assay (Stemina, Todd et al. 2020) needs to be mentioned. This assay uses metabolomics-based key biomarkers to predict developmental toxicity and also displays 85% accuracy (Chang et al. 2022a). In conclusion, despite a lack of OECD guidelines and official validation, a fit-for-purpose toxicodynamics toolbox for animal-free risk assessment of cosmetic ingredients is available and can be put to practical use, providing the possibility, if predicted or needed, to incorporate additional emerging assays.

Practical use of a bioactivity exposure ratio (BER) in animal-free cosmetic ingredient risk assessment

(i) Biological Coverage in Toxicological Risk Assessment

The safety assessment of cosmetic ingredients follows a tiered and iterative process (Dent et al. 2018). Where there is insufficient chemical-specific data to allow a safety decision to be made, techniques such as exposure-based waiving and RAx can be used (Berggren et al. 2017). Where these approaches do not allow a safety conclusion, further data needs to be generated to study the potential biological effects of chemical exposure. Historically in the toxicological risk assessment of cosmetics, biological coverage was provided by performing *in vivo* toxicology studies, usually in one species (the rat) and analysing many different organ systems using a variety of analysis techniques (*e.g.* as described in OECD TG 408). There are clearly strengths and limitations to this approach. The main strengths are that the laboratory rat represents an integrated *in vivo* model, taking into account ADME, and provides an understanding of whether exposure could lead to an adverse effect (a pathological change). The main scientific limitations are that a single strain of laboratory rat is not biologically equivalent to a human, and that whilst analytical tools and our understanding of the molecular basis of disease have developed exponentially in recent years, standardized *in vivo* test guidelines rely on laboratory techniques developed in the nineteenth and twentieth centuries. In NGRA, biological coverage is provided by a suite of mostly human-derived cells or cellular components, and perturbations in the function of such test systems can be analysed at the level of individual genes or the function of individual proteins. This provides unprecedented opportunities to harness the ever-increasing body of biological information now available to develop a more robust and human-relevant approach to assessing the risks posed by chemicals (Krewski et al. 2010), and BERs offer a practical and conservative starting point to tiered decision-making as part of NGRA. A BER is the ratio between a chemical's *in vitro* biological activity-based point of departure and its calculated or measured internal concentration following exposure (Paul-Friedman et al. 2020) and has recently been proposed as a workable solution to allow decision making for complex endpoints such as systemic toxicity without testing on animals (Health Canada 2021; SCCS/1647/2022). Although the BER could be seen as similar to the margin of safety (MoS) traditionally used in cosmetic ingredient safety assessment (SCCS/1647/2022), there are some important differences.

(ii) BER versus MoS

Whilst the MoS uses apical endpoint animal data representing a known adverse health effect as the numerator, the *in vitro* bioactivities used when calculating a BER represent changes in cellular function (*e.g.* protein/receptor interactions, transcriptomic changes), which do not necessarily represent adverse effects. In adverse outcome pathway (AOP) terms these bioactivities are potential molecular initiating or key events, and therefore precede an adverse outcome (Vinken 2013). Secondly, the denominator in MoS calculations ($\text{MoS} = \text{PoD}_{\text{sys}} / \text{systemic exposure dose}$) for cosmetic ingredients.

is depending on an externally applied dose, expressed in mg of the chemical/kg body weight/day (SCCS/1647/2022). BERs on the other hand use internal exposure expressed as a concentration in a body compartment (often blood or plasma) as the denominator (Baltazar et al. 2020). The hypothesis that underpins the use of BERs at the early tier of NGRA is that at exposures which result in internal concentrations which are free of any relevant *in vitro* biological activity, there can be no adversity (Dent et al. 2021). Although the concept of protecting against non-specific effects as opposed to predicting specific toxicities may sound at odds with the traditional animal-based paradigm, in reality this is how the majority of such animal-based safety decisions have been made (Browne et al. 2024). However, transitioning from an integrated 'in vivo biosensor' (*i.e.* an animal test) with a suite of 'in vitro biosensors' clearly requires safety assessors to be confident that all relevant biological activities which could lead to an adverse outcome have been covered and adequately studied *in vitro*. Furthermore, it requires that the dose response for any effects is sufficiently understood to assign a robust PoD.

(iii) Confidence in using BER

If the new approaches are based on perturbations in human-derived cells and evaluate endpoints that are not present in traditional animal toxicology studies, confidence cannot and should not come from validating the results of NAMs against the results of traditional animal tests. Rather, building this confidence requires a new approach (Van der Zalm et al. 2022). One such approach is to benchmark the safety decisions that would be made using NGRA with safety decisions generated by using traditional animal-based approaches. For example, in one evaluation strategy (Middleton et al. 2022), BERs were obtained for known low and high-risk exposures using specific NGRA tools. Those bioactivity tools were an *in vitro* pharmacological profiling panel (Burbank et al. 2024), a cell stress panel (Hatherell et al. 2020) and high throughput transcriptomics (Farmahin et

al. 2017; Reynolds et al. 2020). Used alongside robust PBK modeling, the most sensitive points of departure from this bioactivity toolbox were demonstrated to show an equivalent level of protection to a stand-alone repeat-dose toxicity study in rats for 46 benchmark chemical exposures (Cable et al. accepted 2024). However, because the data analysis techniques used in this evaluation were designed to be conservative (*e.g.* basing the transcriptomics point of departure on a single gene change), the toolbox only correctly identified 7 of the 24 low-risk exposures as low risk. This work illustrates two main points. Firstly, it is important to benchmark the specific NAM toolbox used against relevant known low and high-risk exposures to understand the level of protection and utility it affords. This not only means ensuring the physicochemical properties of the benchmarks are relevant to the chemical being assessed, but also that any lead mode-of-action or toxic effect identified during the assessment is expressed by benchmark chemicals in the dataset. Secondly, because BERs are based on bioactivities that may or may not lead to an adverse outcome, a $BER < 1$ does not necessarily indicate a high risk and may require further data generation, *e.g.* in more physiologically relevant test systems.

Progress in inhalation TTC

(i) Inhalation TTC Development

The use of TTC is described in detail in the SCCS Notes of Guidance, 12th Revision (SCCS/1647/2022). The thresholds proposed are applied for risk assessment of lifetime exposures by the oral route and are not intended to be applied or extended to the inhalation route due to relevant route-specific differences in exposures and associated portal of entry effects in the respiratory tract. However, there is an increased interest in developing TTC for inhalation exposure risk assessments and a need to establish a universal limit for substances with little to no information on their toxicological potency (Bowden et al. 2023). Therefore, establishing inhalation TTC values is important for assessing local respiratory and systemic effects from inhalation exposures to consumer products.

Carthew et al. (2009) published TTC limits based on inhalation exposure-related toxicological studies in rodents. The authors proposed specific limits for systemic and local effects from inhalation exposures with a precise application domain for consumer products or industrial materials, where any inhalation exposure is incidental in most cases. The inhalation TTC dataset consisted of 92 substances with high-quality inhalation exposure studies in rodents and excluded compounds that were genotoxic carcinogens,

in vivo mutagens, heavy metals, dioxins, polychlorinated biphenyls, organophosphates, and polymers. These chemical classes were not considered representative of the ingredients, that are or could be used in aerosols for consumer use. A refined yet conservative approach was later published (Escher et al. 2010) in which no significant difference between local and systemic effects for inhalation TTC values could be observed. Therefore, inhalation TTC limits based on general toxicological effects were recommended. Also, the application domain was less restrictive, which increased the dataset inventory to 203 chemicals. In 2016, the Research Institute for Fragrance Materials (RIFM) initiated a collaborative project with a cosmetic company. The efforts included increasing the chemical inventory, thus expanding the applicability domain beyond the consumer product space and developing a suitable potency-based classification system for analysing local respiratory effects from inhalation exposures. More recently, the inhalation TTC was proposed as a screening tool in a risk-based context for prioritising thousands of chemicals from large inventories (Nelms and Patlewicz 2020). Cramer structural classes were not found to be effective at discriminating against substances based on their potencies. Instead, use of aquatic toxicity modes of action, such as observed in the work of Verhaar et al. (1992), were found to be effective at separating substances in terms of their potencies and deriving inhalation TTC values. The collaboration was expanded with the objective to combine, harmonise, and curate the inhalation exposure datasets, develop an appropriate classification scheme for defining toxicological potency, and derive the 5th percentile inhalation TTC (*inhalTTC*) values for local respiratory and systemic effects of inhalation exposure to substances. The final goal is to publish the outcome and make datasets publicly available.

(ii) Gaps

Knowledge gaps are associated with inhalation exposure-related effects evaluation. It is important to understand the underlying toxicological mechanisms of the adverse effects observed from animal studies, which may help refine the *inhalTTC* values. Since one of the applications of *inhalTTC* is to protect against irritative chemical exposures, it is critical to account for physicochemical properties in addition to structural features of the inhaled substances when developing an appropriate toxicological potency-based classification system. The further development of *inhalTTC* will aid in the safety assessment of inhalation exposures to consumer products and/or their ingredients, and for prioritisation of chemicals for regulatory purposes. However, it is not yet ready for practical application.

Progress in internal TTC (iTTC)

(i) iTTC Development

iTTC refers to a TTC derived from plasma concentration of the chemical substance under evaluation. It is proposed as an evolution of the external dose-based TTC and serves as a promising strategy for advancing non-animal methods, as well as for enhancing the application of TTC in broader contexts. Various groups have made initial attempts to establish an iTTC, each employing unique methodologies. An update on this work was provided in the proceedings of the first workshop (Rogiers et al. 2020). Since then, some progress has been made. Approximately 300 chemicals are included now in the iTTC chemical database and are representative of the TTC chemical space from Munro et al. (1996), Yang et al. (2017) and Patel et al. (2020). To convert the NOAELs for these chemicals into estimates of internal exposure, PBK modeling is employed. This modeling requires chemical-specific input parameters related to ADME. Therefore, a comprehensive literature search was conducted to gather existing *in vivo* pharmacokinetic data and *in vitro* metabolism data. Additionally, *in vitro* Caco-2 permeability and hepatocyte metabolism studies were conducted for about 200 chemicals. The Caco-2 assays measured bidirectional cell permeability, at pHs of 6.5 and 7.4 in the apical and basolateral compartments, respectively, to mimic intestinal conditions. The apparent permeability of the iTTC chemicals measured in Caco-2 cells showed a high variability. The metabolism assays utilised cryopreserved hepatocytes from the species from which the NOAELs were derived, measuring *in vitro* intrinsic clearance through a substrate depletion approach at concentrations of 0.1 and 1 μM . However, metabolism data for about 15% of the assays were deemed unsuitable for clearance determination due to issues such as analytical interference, chemical instability, or rapid loss of test chemical under the assay conditions. PBK model input parameters (*e.g.*, absorption, metabolism) for the iTTC chemicals may come from existing *in vivo* PK data or from *in vitro* studies. As such, the compiled *in vivo* and *in vitro* data on permeability, metabolism and pharmacokinetics were integrated into several kinetic modeling workflows, including top-down, middle-out and bottom-up approaches. These simulations were then compared with corresponding *in vivo* time-course data. For each chemical, the ratios of *in vivo* to predicted values for peak plasma concentration and plasma area under the curve were assessed. Additionally, the performance of specific workflows was evaluated to see whether the predictions exhibited a bias towards over- or under-estimation, and to measure the proximity of the model predictions to the experimental data.

Initial simulations utilising a generic oral rat model with *in vivo* inputs (top-down) demonstrated good concordance with *in vivo* oral rat pharmacokinetic data, while those using *in vitro* inputs (bottom-up) yielded more variable results in terms of concordance.

(ii) Gaps

Areas for improvement in the bottom-up modeling workflows have been identified and are the primary focus of the ongoing work. Case studies highlighting the practical applications of iTTC in human safety assessment will continue to be pursued. In the meantime, a conservative interim iTTC of 1 μM has been proposed.

OECD guidance for the use of QSARs in regulatory assessments

The OECD definition of NAMs includes *in chemico*, *in silico*, *in vitro* and *in vivo* approaches that are aligned with the '3R' principles and bridge the gaps in moving towards animal-free regulatory assessments. The use of NAMs is intended to increase the throughput of chemicals assessments, while using the best available and most relevant science, and simultaneously reducing the reliance on animal testing. However, it should be noted that the use of *in vivo* data represents an important inhibitory factor for the safety assessment of cosmetics in Europe as only animal-free NAMs are allowed.

The QSAR Toolbox (www.qsartoolbox.org), initially launched in 2008 with the European Chemicals Agency (ECHA) and rapidly joined as a co-lead, was initially foreseen as a software application to support grouping and RAX approaches. While it has been and continues to be used for this purpose, the Toolbox also includes a repository of experimental data on more than 100,000 chemicals and three million datapoints. Users can retrieve experimental data, physical–chemical properties, and simulate metabolism and degradation products from parent chemicals. In addition, the QSAR Toolbox Repository (<https://repository.qsartoolbox.org/>) houses a variety of third-party software applications that can support computational predictions of chemical effects. The software is updated regularly to expand the functionality and improve the user interface, with new versions released on an approximately annual basis.

In addition to the software application itself, the OECD published principles for the validation of QSARs (OECD 2004) models and guidance for the application of validation principles to QSAR models (OECD 2007) in a regulatory context to help assure the reliability and robust use of these approaches in chemical risk assessment. The principles for

QSAR validation were directly adapted from the principles for validation of OECD test guidelines for using experimental approaches for evaluating chemical hazards (OECD 2005). Along with guidance for evaluating the scientific validity of QSARs, templates for providing information on models (QSAR Model Reporting Format (QMRF)) and predictions (QSAR Prediction Reporting Format (QPRF)) were developed to ensure necessary information for the evaluation was reported in a standardised way.

In response to the identified need to build on the established OECD principles for evaluating QSAR models, a QSAR Assessment Framework (QAF) was developed for systematic and harmonised guidance for the regulatory assessment of QSAR models and predictions (also from multiple models). The guidance was primarily intended for regulatory authorities; however, the same principles could be used as a reference for all stakeholders developing and using QSARs for regulatory purposes. An updated QAF was published in 2024 (OECD 2024) and includes a new QSAR Results Reporting Format (QRRF) developed for standardising the reporting of results from multiple QSAR predictions. Important to know is that the guidance, various reporting formats, checklists for evaluating criteria associated with each assessment element, and examples for applying the QAF in real-world, regulatory examples are provided on the OECD website (https://www.oecd.org/en/publication/s/2023/11/q-sar-assessment-framework-guidance-for-the-regulatory-assessment-of-quantitative-structure-activity-relationship-models-and-predictions_9b064821.html).

Discussion and conclusion

At the second SCCS Workshop on ‘*The way forward for assessing the human health safety of cosmetics in the EU*’, experts with knowledge and experience of applying NAMs and NGRA as the key methodologies for animal-free safety assessment of cosmetics and their ingredients were brought together. In preparation of the one-day workshop, specific questions were sent to the speakers beforehand to elicit science-based and reliable responses on a range of complex issues highlighted as potential gaps in animal-free risk assessment of chemical substances used in cosmetic ingredients.

One of the key questions concerned the criteria necessary for the acceptance of PBK models when assessing a new chemical (*e.g.* cosmetic ingredients) for which no in vivo data exists. This scenario arises when human exposure is limited—and therefore relevant human data is not available—and animal testing is prohibited under the Cosmetics Regulation. In such cases, where an ab initio NGRA for systemic toxicity is required, a defined set of criteria needs to be

met. Development of such decision factors is crucial. Clear and well-supported suggestions were discussed, drawing on considerations previously explored by EFSA. It was emphasised that PBK models can provide valuable insights into the biokinetics and toxicity of new chemicals, provided that key elements are carefully addressed, namely: the quality of the input data, the appropriateness of the model structure, calibration processes, sensitivity analysis, and alignment with the regulatory requirements. However, in the absence of in vivo data, the acceptance of PBK models will need rigorous validation/evaluation, transparent model documentation, and the use of advanced computational tools to address uncertainties inherent to the model simulations. Finally, the workshop underscored the importance of continued dialogue between model developers and risk assessors to enhance effective PBK model application for the assessment of chemicals, in this case cosmetic ingredients.

Further insight into the increasing importance of PBK use emerged from the presentation on (Q)IVIVE. Based on practical case studies involving cosmetic ingredients, it was demonstrated that (Q)IVIVE serves as a critical tool for enabling the implementation of NAMs and NGRA in cosmetic safety assessments. Indeed, (Q)IVIVE allows for the prediction of internal concentrations over time in humans (or animals), facilitating the identification of PoDs for hazard characterisation and the determination of safe use concentrations for cosmetic ingredients. As this methodology continues to evolve, it is expected to increasingly play a central role in regulatory decision-making, enabling a shift toward more protective, human-relevant, and non-animal-based testing frameworks.

The same case studies were used to highlight the value of developing a comprehensive toxicodynamics toolbox, containing a variety of different animal-free methodologies. In particular, the lack of validated animal-free assays for developmental/reproductive toxicity remains a critical gap, especially given the relevance of these endpoints in standard risk assessments of new substances, such as preservatives.

A newly developed template outlining readiness criteria for NAMs was introduced by the EU’s Joint Research Centre at the Workshop. These criteria have to be fulfilled before a method can proceed to either validation, transferability or peer-review. It was suggested that they could support the SCCS in evaluating the scientific validity of non-animal, non-validated methods before their potential use in the safety assessment of cosmetic ingredients.

During the discussion on the use of the BER as a practical and robust basis for safety decisions, it was noted that more confidence in the approach is needed. To build this trust and to better understand the level of protection provided by NGRA application, it is important to evaluate not only low-risk compounds but also high-risk substances. Furthermore,

there is a need to establish a flexible and practical range of BER thresholds, adaptable to the characteristics of different substances and the available supporting data.

Further progress has been made with respect to the application of TTC for inhalation exposure as an effective animal-alternative tool for the safety assessment of data-poor substances. However, it is not yet sufficiently developed for use in a regulatory context. Similarly, some advancements in iTTC have been made, but further work is needed. Specific areas for improvement in the bottom-up modeling workflows have been identified and are the focus of ongoing work. Once finalised, the outcomes will be openly shared and documented to support broader implementation.

The work of the OECD in standardising new technologies is highly valued. However, the OECD's current definition of NAMs, which allows for the inclusion of new animal data, is not compatible with the regulatory framework for cosmetics. This represents a need for harmonising the terminology.

Finally, several issues were raised during the workshop that merit further discussion and elaboration:

***Regulatory confidence in NAMs:** A central question remains whether NAMs provide the same level of confidence to regulators as traditional *in vivo* methodology. Encouragingly, for certain toxicological endpoints there seems to be a clear shift from 'unlikely' to 'likely', but challenges remain for more complex endpoints.

***Challenges from the animal testing ban:** The ban of animal testing in the cosmetic sector brings significant challenges for both the development of new ingredients/products as well as for regulatory risk assessors. While several promising tools are available, several key endpoints are not covered and a clear harmonised guidance is missing.

***Alignment of NGRA with regulatory needs:** The NGRA approach is a valuable alternative that can be used in risk assessment, but its practical application requires better alignment with the needs of assessors/regulators. Furthermore, most NGRA case studies of cosmetic ingredients that are available in the open literature follow a similar set of methods to retrieve valuable data. Higher methodologic variation, possibly introduced by different stakeholders, could further contribute to building trust.

***Scientific validity versus validation and weight of evidence:** For concepts such as 'scientifically valid' and 'qualification' (new terminology being developed in an EFSA project for non-validated nanomaterials in food and feed), a set of criteria must be further developed as they will provide a pragmatic solution to make further progress in the practical application of animal-free methodology. Indeed, they can represent a realistic step forward in the complex validation process. Risk assessors and regulatory bodies also need to gain experience and expertise in the effective

application of a 'WoE' approach to make reliable safety assessment decisions. It remains unclear, however, who finally determines the acceptance of innovative methodology such as NGRA in risk assessment of cosmetics and their ingredients.

***Quantitative versus hazard-based assessment:** In Europe, cosmetics and their ingredients are safe due to the stringent Cosmetic Regulation based on quantitative risk assessment. The risk assessment process is science-based and rigorously described in the SCCS Notes of Guidance, which are regularly updated and used worldwide as a reference. A shift toward purely hazard-based decision-making will clearly undermine the scientific rigor and trust in the risk assessment process.

*** Additional questions,** although not directly discussed in this workshop, have been identified for future exploration:

- Can a mode of action (if any) be reliably derived through RAX approaches?
- While a systemic and harmonised NGRA framework for cosmetic ingredients/chemicals is supported by case studies, how can nano-size substances, mixtures and aggregate exposure be addressed?
- When NGRA is applied for systemic toxicity, are all potential toxicodynamic effects adequately covered?
- When the BER concept is used, is the distinction between biological activity and adversity appropriately considered in decision making?

In conclusion, to advance the use of NGRA as an animal-free risk assessment strategy for both new and existing cosmetic ingredients, a close interaction between different stakeholders in the field is essential. Such cooperation is key to fostering a shared understanding and promoting regulatory acceptance of NAM-based safety assessments.

Acknowledgements The contribution of all participants to the workshop was greatly appreciated. And a special thanks for the excellent organisation of the Workshop goes to the SCCS Secretariat.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Alexander-White C, Bury D, Cronin M, Dent M, Hack E, Hewitt NF, Kenna G, Naciff J, Ouedraogo G, Schepky A, Mahony C, Cosmetics Europe (2022) A 10-step framework for use of read-across (RAX) in next generation risk assessment (NGRA) for cosmetics safety assessment. *Regul Toxicol Pharmacol* 129:105094. <https://doi.org/10.1016/j.yrtph.2021.105094>
- Andersen ME (1995) Development of physiologically based pharmacokinetic and physiologically based pharmacodynamic models for applications in toxicology and risk assessment. *Toxicol Lett* 79(1–3):35–44. [https://doi.org/10.1016/0378-4274\(95\)03355-0](https://doi.org/10.1016/0378-4274(95)03355-0)
- Bal-Price A, Hogberg HT, Crofton KM, Daneshian M, FitzGerald RE, Fritsche E, Heinonen T, Hougaard Bennekou S, Klima S, Piersma AH, Sachana M, Shafer TJ, Terron A, Monnet-Tschudi F, Viviani B, Waldmann T, Westerink RHS, Wilks MF, Witters H, Zurich MG, Leist M (2018) Recommendation on test readiness criteria for new approach methods in toxicology: exemplified for developmental neurotoxicity. *Altex* 35(3):306–352. <https://doi.org/10.14573/altex.1712081>
- Baltazar M, Cable S, Carmichael P, Cubberley R, Cull T, Delagrangé M, Dent M, Hatherell S, Houghton J, Kukic P, Li H, Lee M-Y, Malcomber S, Middelton A, Moxon T, Nathanail A, Nicol B, Pendlington R, Reynolds G, Reynolds J, White A, Westmoreland C (2020) A next-generation risk assessment case study for coumarin in cosmetic products. *Toxicol Sci* 176:236–252. <https://doi.org/10.1093/toxsci/kfaa048>
- Bell SM, Chang X, Wambaugh JF, Allen DG, Bartels M, Brouwer KLR, Casey WM, Choski N, Ferguson SS, Fraczkiwicz G, Jarabek AM, Ke A, Lumen A, Lynn SG, Paini A, Price PS, Ring C, Simon TW, Sipes NS, Sprankle CS, Strickland J, Troutman J, Wetmore BA, Kleinstreuer NC (2018) *In vitro* to *in vivo* extrapolation for high throughput prioritization and decision making. *Toxicol in Vitro* 47:213–227. <https://doi.org/10.1016/j.tiv.2017.11.016>
- Berggren E, White A, Ouedraogo G, Paini A, Richarz A, Bois F, Exner T, Leite S, Van Grunsven L, Worth A, Mahony C (2017) *Ab initio* chemical safety assessment: a workflow based on exposure considerations and non-animal methods. *Comput Toxicol* 4:31–44. <https://doi.org/10.1016/j.comtox.2017.10.001>
- Bowden AM, Escher SE, Rose J, Sadekar N, Patlewicz G, O’Keeffe L, Bury D, Hewitt NJ, Giusti A, Rothe H (2023) Workshop report: challenges faced in developing inhalation thresholds of Toxicological Concern (TTC) - state of the science and next steps. *Regul Toxicol Pharmacol* 142:105434. <https://doi.org/10.1016/j.yrtph.2023.105434>
- Breen M, Ring CL, Kreutz A, Goldsmith M-R, Wambaugh JF (2021) High-throughput PBTK models for *in vitro* to *in vivo* extrapolation. *Expert Opin Drug Metab Toxicol* 17(8):903–921. <https://doi.org/10.1080/17425255.2021.1935867>
- Browne P, Paul Friedman K, Boekelheide K, Thomas R (2024) Adverse effects in traditional and alternative toxicity tests. *Regul Toxicol Pharmacol* 148:105579. <https://doi.org/10.1016/j.yrtph.2024.105579>
- Burbank M, Kukic P, Ouedraogo G, Kenna JG, Hewitt NJ, Armstrong D, Otto-Bruc A, Ebmeyer J, Boettcher M, Willox I, Mahony C (2024) *In vitro* pharmacologic profiling aids systemic toxicity assessment of chemicals. *Toxicol Appl Pharmacol* 492:117131. <https://doi.org/10.1016/j.taap.2024.117131>
- Cable S, Baltazar M, Bunglawala F, Carmichael P, Contreas L, Dent M, Houghton J, Kukic P, Malcomber S, Nicol B, Przybylak KR, Punt A, Reynolds G, Reynolds J, Scott S, Tang D, Middleton A (2024) Advancing systemic toxicity risk assessment: evaluation of a NAM-based toolbox approach. *Toxicol Sci* 204(1):79–95. <https://doi.org/10.1093/toxsci/kfae159>
- Carthew P, Clapp C, Gutsell S (2009) Exposure based waiving: the application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. *Food Chem Toxicol* 47(6):1287–1295. <https://doi.org/10.1016/j.foct.2009.02.024>
- Chang X, Palmer J, Lumen A, Lee UJ, Ceger P, Mansouri K, Sprankle C, Donley E, Bell S, Knudsen TB, Wambaugh J, Cook B, Allen D, Kleinstreuer N (2022a) Quantitative *in vitro* to *in vivo* extrapolation for developmental toxicity potency of valproic acid analogues. *Birth Defects Res* 114(16):1037–1055. <https://doi.org/10.1002/bdr2.2019>
- Chang X, Tan YM, Allen DG, Bell S, Brown PC, Browning L, Ceger P, Gearhart J, Hakkinen P, Kabadi SV, Kleinstreuer N, Lumen N, Matheson J, Paini A, Pangburn HA, Petersen EJ, Reinke EN, Ribeiro AJS, Sipes N, Sweeney LM, Wambaugh JF, Wange R, Wetmore BA, Mumtaz M (2022b) IVIVE: facilitating the use of *in vitro* toxicity data in risk assessment and decision making. *Toxics* 10(5):232. <https://doi.org/10.3390/toxics10050232>
- Cook C, Albrecht L, Cary G, et al (2022) EPA’s concentration versus time database: a resource for extrapolating toxicokinetic trends across chemicals. Poster presented at the SOT Annual Meeting in San Diego. https://epa.figshare.com/articles/poster/EPA_s_Concentration_versus_Time_Database_A_Resource_for_Extrapolating_Toxicokinetic_Trends_Across_Chemicals/18624749?file=34335176. Accessed 28 October 2025
- Crouzet T, Grignard E, Brion F, Blanc EB, Podechard N, Langouet S, Alonso-Magdalena P, Hubert P, Kim MJ, Audouze K (2023) ReadEDTest: a tool to assess the readiness of *in vitro* test methods under development for identifying endocrine disruptors. *Environ Int* 174:107910. <https://doi.org/10.1016/j.envint.2023.107910>
- Dent M, Vaillancourt E, Thomas R, Carmichael P, Ouedraogo G, Kojima H, Barroso J, Ansell J, Barton-Maclaren TS, Bennekou SH, Boekelheide K, Ezezdandam J, Field J, Fitzpatrick S, Hatao M, Kreiling R, Lorencini M, Mahoni C, Montemayor B, Mazarro-Costa R, Oliveira J, Rogiers V, Smegal D, Taalman R, Tokura Y, Verma R, Wilett C, Yang C (2021) Paving the way for application of next generation risk assessment to safety decision-making for cosmetic ingredients. *Regul Toxicol Pharmacol* 125:105026. <https://doi.org/10.1016/j.yrtph.2021.105026>
- Dent M, Teixeira Amaral R, Amores Da Silva P, Ansell J, Boislevé F, Hatao M, Hirose A, Kasai Y, Kern P, Kreiling R, Milstein S, Montemayor B, Oliveira J, Richarz A, Taalman R, Vaillancourt E, Verma R, Vieira O’Reilly Cabral Posada N, Weiss C, Kojima H (2018) Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. *Comput Toxicol* 7:20–26. <https://doi.org/10.1016/j.comtox.2018.06.001>
- Ebmeyer J, Najjar A, Lange D, Boettcher M, Voss S, Brandmair K, Meinhardt J, Kuehl J, Hewitt NJ, Krueger CT, Schepky A (2024) Next generation risk assessment: an *ab initio* case study to assess the systemic safety of the cosmetic ingredient, benzyl salicylate, after dermal exposure. *Front Pharmacol* 15:1345992. <https://doi.org/10.3389/fphar.2024.1345992>
- Escher SE, Tluczkiwicz I, Batke M, Bitsch A, Melber C, Kroese ED, Buist HE, Mangelsdorf I (2010) Evaluation of inhalation TTC values with the database RepDose. *Regul Toxicol Pharmacol* 58:259–274. <https://doi.org/10.1016/j.yrtph.2010.06.009>
- EU Reference Laboratory for alternatives to animal testing (EURL ECVAM), EURL ECVAM test method submission. https://joint-research-centre.ec.europa.eu/reference-measurement/european-un-ion-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam/alternative-methods-toxicity-testing/validation-and-submission-process/eurl-ecvam-test-method-submission_en. Accessed 27 October 2025
- European Commission Directorate-General for Environment (2025) Conference report – Brussels 25 October 2024 – 2nd Commission conference on the Roadmap Towards Phasing out

- Animal Testing for Chemical Safety Assessments, Publications Office of the European Union. <https://data.europa.eu/doi/https://doi.org/10.2779/5469181>. Accessed 27 October 2025
- European Commission EC (2009) Regulation No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products (recast) - Current consolidated version: 01/05/2025. OJ L 342: 59–209. <http://data.europa.eu/eli/reg/2009/1223/oj>. Accessed 27 October 2025
- European Medicine Agency (EMA) (2018) Guideline on the Qualification and Reporting of Physiologically Based Pharmacokinetic (PBPK) Modelling and simulation. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpbk-modelling-and-simulation_en.pdf. Accessed 27 October 2025
- Farmahin R, Williams A, Kuo B, Cheplev N, Thomas R, Barton-Maclaren T, Curran IH, Nong A, Wade MG, Yauk C (2017) Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. *Arch Toxicol* 91:2045–2065. <https://doi.org/10.1007/s00204-016-1886-5>
- Géniès C, Jeanjean C, Najjar A et al (2025a) Characterization of the *in vitro* penetration and first-pass metabolism of genistein and daidzein using human and pig skin explants and Phenion full-thickness skin models. *J Appl Toxicol* 45(2):200–209. <https://doi.org/10.1002/jat.4689>
- Géniès C, Jeanjean C, Najjar A et al (2025b) Effect of vehicle on the *in vitro* penetration and metabolism of genistein and daidzein in *ex vivo* skin explants and the Phenion full-thickness skin model. *J Appl Toxicol* 45(2):210–221. <https://doi.org/10.1002/jat.4693>
- Gilmour N, Alépée N, Hoffmann S, Kern PS, van Vliet E, Bury D et al (2023) Applying a next generation risk assessment framework for skin sensitisation to inconsistent new approach methodology information. *ALTEX* 40:439–451. <https://doi.org/10.14573/altex.2211161>
- Hatherell S, Baltazar M, Reynolds J, Carmichael P, Dent M, Li H et al (2020) Identifying and characterizing stress pathways of concern for consumer safety in next-generation risk assessment. *Toxicol Sci* 176:11–33. <https://doi.org/10.1093/toxsci/kfaa054>
- Health Canada (2021) Science approach document - Bioactivity exposure ratio: Application in priority setting and risk assessment. <https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/science-approach-document-bioactivity-exposure-ratio-application-priority-setting-risk-assessment.html>. Accessed 27 October 2025
- Holzer AK, Dresser N, Pallocca G, Mangerich A, Stacey G, Dipalo M, Van de Water B, Rovida C, Wirtz PH, Van Vugt B, Panzarella G, Hartung T, Terron A, Mangas I, Herzler M, Marx-Stoelting P, Coecke S, Leist M (2023) Acceptance criteria for new approach methods in toxicology and human health-relevant life science research—Part I. *ALTEX* 40(4):706–712. <https://doi.org/10.14573/altex.2310021>
- Jamalpoor A, Hartvelt S, Dimopoulou M, Zwetsloot T, Brandsma I, Racz PI, Osterlund T, Hendriks G (2022) A novel human stem cell-based biomarker assay for *in vitro* assessment of developmental toxicity. *Birth Defects Res* 114(19):1210–1228. <https://doi.org/10.1002/bdr2.2001>
- Katsanou ES, Spyropoulou A, Batakis P, Kyriakopoulou K, Repouskou A, Spilioti E, Vogiatzi E, Mouzaki-Paxinou AC, Busquet F, Audouze K, Blanc EB, Kim MJ, Grignard E, Hubert P, Langezaal I, Munn S, Machera K (2025) Methodology for the identification and selection of Endocrine Disruptor-relevant methods suitable for validation: A PEPPER project. Ready to be submitted to *Front Toxicol*
- Krebs A, Waldmann T, Wilks M F, Van Vugt-Lussenburg B M A, Van der Burg B, Terron A, Steger-Hartmann T, Ruegg J, Rovida C, Pedersen E, Pallocca G, Luijten M, Leite S B, Kustermann S, Kamp H, Hoeng J, Hewitt P, Herzler M, Hengstler JG, Heinonen T et al (2019) Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data. *ALTEX* 36(4):682–699. <https://doi.org/10.14573/altex.1909271>
- Krewski D, Acosta D, Andersen M, Anderson H, Bailar JC, Boekelheide K, Staff of Committee on Toxicity Testing and Assessment of Environmental Agents (2010) Toxicity testing in the 21st century: a vision and a strategy. *J Toxicol Environ Health B Crit Rev* 13(2–4):51–138. <https://doi.org/10.1080/10937404.2010.483176>
- Kroes R, Renwick AG, Cheeseman M, Kleiner J, Mangelsdorf I, Piersma A, Schilter B, Schlatter J, van Schothorst F, Vos JG, Würtzen G, European branch of the International Life Sciences Institute (2004) Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food Chem Toxicol* 42(1):65–83. <https://doi.org/10.1016/j.fct.2003.08.006>
- Kuepfer L, Niederalt C, Wendl T et al (2016) Applied concepts in PBPK modeling: How to build a PBPK/PD model? *CPT Pharmacometrics Syst Pharmacol* 5(10):516–531. <https://doi.org/10.1002/psp4.12134>
- Kulkarni PR, Youssef AS, Argikar AA (2021) Prediction of drug clearance from enzyme and transporter kinetics. *Methods Mol Biol* 2342:369–417. https://doi.org/10.1007/978-1-0716-1554-6_14
- Louisse J, Beekmann K, Rietjens IMCM (2017) Use of physiologically based kinetic modeling-based reverse dosimetry to predict *in vivo* toxicity from *in vitro* data. *Chem Res Toxicol* 30(1):114–125. <https://doi.org/10.1021/acs.chemrestox.6b00302>
- Madden JC, Pawar G, Cronin MTD, Webb S, Tan Y-M, Paini A (2019) *In silico* resources to assist in the development and evaluation of physiologically-based kinetic models. *Comput Toxicol* 11:33–49. <https://doi.org/10.1016/j.comtox.2019.03.001>
- Middleton A, Reynolds J, Cable C, Baltazar M, Li H, Bevan S et al (2022) Are non-animal systemic safety assessments protective? A toolbox and workflow. *Toxicol Sci* 189(1):124–147. <https://doi.org/10.1093/toxsci/kfac068>
- Munro IC, Ford RA, Kennepohl E, Sprenger JG (1996) Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food Chem Toxicol* 34(9):829–867. [https://doi.org/10.1016/s0278-6915\(96\)00049-x](https://doi.org/10.1016/s0278-6915(96)00049-x)
- Naga D, Parrott N, Ecker GF, Olivares-Morales A (2022) Evaluation of the success of high-throughput physiologically based pharmacokinetic (HT-PBPK) modeling predictions to inform early drug discovery. *Mol Pharm* 19(7):2203–2216. <https://doi.org/10.1021/acs.molpharmaceut.2c00040>
- Najjar A, Punt A, Wambaugh J, Paini A, Ellison C, Fragki S, Bianchi E, Zhang F, Westerhout J, Mueller D, Li H, Shi Q, Gant TW, Botham P, Bars R, Piersma A, van Ravenzwaay B, Kramer N (2022) Towards best use and regulatory acceptance of generic physiologically based kinetic (PBK) models for *in vitro*-to-*in vivo* extrapolation (IVIVE) in chemical risk assessment. *Arch Toxicol* 96(12):3407–3419. <https://doi.org/10.1007/s00204-022-03356-5>
- Najjar A, Gregoire S, Nicol B et al (2024a) *In vitro* to *in vivo* extrapolation to derive a metabolism factor for estimating the aggregate exposure to salicylic acid after dermal exposure of its esters. *Arch Toxicol* 98(7):2199–2211. <https://doi.org/10.1007/s00204-024-03749-8>
- Najjar A, Hamadeh A, Krause S, Schepky A, Edginton A (2024b) Global sensitivity analysis of Open Systems Pharmacology Suite physiologically based pharmacokinetic models. *CPT Pharmacometrics Syst Pharmacol* 13(12):2052–2067. <https://doi.org/10.1002/psp4.13256>
- Najjar A, Kuhl J, Lange D et al (2024c) Next-generation risk assessment read-across case study: application of a 10-step framework to derive a safe concentration of daidzein in a body lotion. *Front*

- Pharmacol 15:1421601. <https://doi.org/10.3389/fphar.2024.1421601>
- Najjar A, Lange D, Génies C et al (2024d) Development and validation of PBPK models for genistein and daidzein for use in a next-generation risk assessment. *Front Pharmacol* 15:1421650. <https://doi.org/10.3389/fphar.2024.1421650>
- Najjar A, Ellison CA, Gregoire S, Hewitt NJ (2023) Practical application of the interim internal threshold of toxicological concern (iTTC): a case study based on clinical data. *Arch Toxicol* 97(1):155–164. 10.1007/s00204-022-03371-6. Erratum in *Arch Toxicol* 97(1):311. <https://doi.org/10.1007/s00204-022-03398-9>
- Nelms MD, Patlewicz G (2020) Derivation of new threshold of toxicological concern values for exposure via inhalation for environmentally-relevant chemicals. *Front Toxicol* 2:580347. <https://doi.org/10.3389/ftox.2020.580347>
- OECD (2004) OECD Principles for the Validation, for Regulatory Purposes, of (Q)SAR Models, No. 49. OECD Publishing, Paris. <http://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/assessment-of-chemicals/oecd-principles-for-the-validation-for-regulatory-purposes-of-quantitative-structure-activity-relationships-p-models.pdf>. Accessed 27 October 2024
- OECD (2005) Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, OECD Series on Testing and Assessment, No. 34. OECD Publishing, Paris. <https://doi.org/10.1787/e1f1244b-en>. Accessed 28 October 2025
- OECD (2007) Guidance document on the validation of (Quantitative) structure-activity relationship [(Q)SAR] Models. ENV/JM/MONO(2007)2, No. 69, OECD Publishing, Paris. <https://doi.org/10.1787/9789264085442-en>
- OECD (2014) Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. No. 69. OECD Publishing, Paris. <https://doi.org/10.1787/9789264085442-en>. Accessed 27 October 2025
- OECD (2018) Guidance Document on Good *In vitro* Method Practices (GIVIMP), OECD Series on Testing and Assessment, No. 286. OECD Publishing, Paris. <https://doi.org/10.1787/9789264304796-en>. Accessed 28 October 2025
- OECD (2021) Guidance document on the characterisation, validation and reporting of Physiologically Based Kinetic (PBK) models for regulatory purposes, OECD Series on Testing and Assessment, No. 331. OECD Publishing, Paris. <https://doi.org/10.1787/d0de241f-en>. Accessed 28 October 2025
- OECD (2023) (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions, OECD Series on Testing and Assessment, No. 386. OECD Publishing, Paris. <https://doi.org/10.1787/d96118f6-en>. Accessed 28 October 2025
- OECD (2024) (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions, Second Edition, OECD Series on Testing and Assessment, No. 405. OECD Publishing, Paris. <https://doi.org/10.1787/bbdac345-en>. Accessed 28 October 2025
- OECD (2025) Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents, OECD Guidelines for the Testing of Chemicals Section 4. OECD Publishing, Paris. <https://doi.org/10.1787/9789264070707-en>. Accessed 28 October 2025
- Paini A, Tan YM, Sachana M, Worth A (2021) Gaining acceptance in next generation PBK modelling approaches for regulatory assessments - An OECD international effort. *Comput Toxicol* 18:100163. <https://doi.org/10.1016/j.comtox.2021.100163>
- Patel A, Joshi K, Rose J, Laufersweiler M, Felter SP, Api AM (2020) Bolstering the existing database supporting the non-cancer Threshold of Toxicological Concern values with toxicity data on fragrance-related materials. *Regul Toxicol Pharmacol* 116:104718. <https://doi.org/10.1016/j.yrtph.2020.104718>
- Paul-Friedman K, Gagne M, Loo LH, Karamertzanis P, Netzeva T, Sobanski T et al (2020) Utility of *in vitro* bioactivity as a lower bound estimate of *in vivo* adverse effect levels and in risk-based prioritization. *Toxicol Sci* 173:202–225. <https://doi.org/10.1093/toxsci/kfz201>
- Pearce RG, Setzer RW, Strobe CL, Wambaugh JF, Sipes NS (2017) Httk: R package for high-throughput toxicokinetics. *J Stat Softw* 79(4):1–26. <https://doi.org/10.18637/jss.v079.i04>
- Petersen EJ, Elliott JT, Gordon J, Kleinstreuer NC, Reinke E, Roesslein M, Toman B (2023) Technical framework for enabling high quality measurements in new approach methodologies (NAMs). *Altex* 40(1):174–186. <https://doi.org/10.14573/altex.2205081>
- Pletz J, Blakeman S, Paini A, Parissis N, Worth A, Andersson AM, Frederiksen H, Sakhi AK, Thomsen C, Bopp SK (2020) Physiologically based kinetic (PBK) modelling and human biomonitoring data for mixture risk assessment. *Environ Int* 143:105978. <https://doi.org/10.1016/j.envint.2020.105978>
- Punt A, Louisse J, Beekmann K, et al (2022) Predictive performance of next generation human physiologically based kinetic (PBK) model predictions based on *in vitro* and *in silico* input data. *ALTEX* 39(2):221–234. <https://doi.org/10.14573/altex.2108301>
- Reynolds J, Malcomber S, White A (2020) A Bayesian approach for inferring global points of departure from transcriptomics data. *Comput Toxicol* 16:100138. <https://doi.org/10.1016/j.comtox.2020.100138>
- Rogiers V, Benfenati E, Bernauer U, Bodin L, Carmichael P, Chaudhry Q, Coenraads PJ, Cronin MTD, Dent M, Dusinska M, Ellison C, Ezendam J, Gaffet E, Galli CL, Goebel C, Granum B, Hollnagel HM, Kern PS, Kosemund-Meynen K, Ouédraogo G, Panteri E, Rousselle C, Stepnik M, Vanhaecke T, von Goetz N, Worth A (2020) The way forward for assessing the human health safety of cosmetics in the EU - Workshop proceedings. *Toxicology* 436:152421. <https://doi.org/10.1016/j.tox.2020.152421>
- SCCS (2022) Scientific opinion on Genistein and Daidzein SCCS/1641/2022 final version. https://health.ec.europa.eu/system/files/2023-08/sccs_o_263.pdf. Accessed 28 October 2025
- SCCS (2023) The SCCS Notes of Guidance for the testing of cosmetic Ingredients and their safety evaluation SCCS/1647/22, 12th revision, 15 May 2023, corrigendum 1 on 26 October 2023, corrigendum 2 on 21 December 2023. https://health.ec.europa.eu/document/download/32a999f7-d820-496a-b659-d8c296cc99c1_en?file_name=sccs_o_273_final.pdf. Accessed 28 October 2025
- SCCS (2024) Scientific opinion on Benzophenone-4 SCCS/1660/2023. https://health.ec.europa.eu/latest-updates/sccs-final-opinion-benzophenone-4-2024-04-04_en. Accessed 28 October 2025
- Sepelri S, De Win D, Heymans A, Van Goethem F, Rodrigues RM, Rogiers V, Vanhaecke T (2025) Next generation risk assessment of hair dye HC yellow no. 13: ensuring protection from liver steatogenic effects. *Regul Toxicol Pharmacol* 159:1–14. <https://doi.org/10.1016/j.yrtph.2025.105794>
- Tan Y-M, Chan M, Chukwudebe A et al (2020) PBPK model reporting template for chemical risk assessment applications. *Regul Toxicol Pharmacol* 115:10469. <https://doi.org/10.1016/j.yrtph.2020.104691>
- Thompson CV, Firman JW, Goldsmith MR et al (2021) A systematic review of published physiologically-based kinetic models and an assessment of their chemical space coverage. *Altern Lab Anim* 49(5):197–208. <https://doi.org/10.1177/02611929211060264>
- Todd JZ, Saili KS, Rush N, Kothiyi P, Judson RS, Houck KA, Hunter ES, Baker NC, Palmer JA, Thomas RS, Knudsen TB (2020) Profiling the ToxCast library with a pluripotent human (H9) stem cell line-based biomarker assay for developmental toxicity. *Toxicol Sci* 174(2):189–209. <https://doi.org/10.1093/toxsci/kfaa014>
- U.S. Food and Drug Administration (US FDA) (2018) Physiologically Based Pharmacokinetic Analyses — Format and Content Guidance for Industry. <https://www.fda.gov/files/drugs/published/P>

- [physiologically-Based-Pharmacokinetic-Analyses-%E2%80%944-Format-and-Content-Guidance-for-Industry.pdf](#). Accessed 27 October 2025
- Van der Zalm AJ, Barroso J, Browne P, Casey W, Gordon J, Henry TR, Kleinstreuer NC, Lowit AB, Perron M, Clippinger AJ (2022) A framework for establishing scientific confidence in new approach methodologies. *Arch Toxicol* 96(11):2865–2879. <https://doi.org/10.1007/s00204-022-03365-4>
- Verhaar HJM, van Leeuwen CJ, Hermens JLM (1992) Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere* 25:471–491. [https://doi.org/10.1016/0045-6535\(92\)90280-5](https://doi.org/10.1016/0045-6535(92)90280-5)
- Vinken M (2013) The adverse outcome pathway concept: a pragmatic tool in toxicology. *Toxicology* 312:158–165. <https://doi.org/10.1016/j.tox.2013.08.011>
- Walder L, Pallocca G, Bastos LF, Beekhuijzen M, Busquet F, Constantino H, Corvaro M, Courtot L, Escher B, Fernandez R, Gougeon E, Hansell L, Herzler M, Holden L, Hornek-Gausterer R, Irizar A, Kandarova H, Kern P, Kolle S, Lacasse K, Lee I, Macmillan DS, Maxwell G, Moriarty O, Nadzialek S, Pochat J, Reid K, Revel M, Ritskes-Hoitinga M, Sobanski T, Stoddart G, Underhill D, Veillette M, Vriend J, Westmoreland C, Baines J (2025) EU roadmap for phasing out animal testing for chemical safety assessments: recommendations from a multi-stakeholder roundtable. *Altex* 42(3):435–450. <https://doi.org/10.14573/altex.2503241>
- Wambaugh JF, Hughes MF, Ring CL et al (2018) Evaluating *in vitro-in vivo* extrapolation of toxicokinetics. *Toxicol Sci* 163(1):152–169. <https://doi.org/10.1093/toxsci/kfy020>
- Wetmore BA, Wambaugh JF, Ferguson SS et al (2012) Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* 125(1):157–174. <http://doi.org/10.1093/toxsci/kfy020>
- Wetmore BA, Wambaugh JF, Ferguson SS et al (2013) Relative impact of incorporating pharmacokinetics on predicting *in vivo* hazard and mode of action from high-throughput *in vitro* toxicity assays. *Toxicol Sci* 132(2):327–346. <https://doi.org/10.1093/toxsci/kft012>
- Wetmore BA, Wambaugh JF, Allen B et al (2015) Incorporating high-throughput exposure predictions with dosimetry-adjusted *in vitro* bioactivity to inform chemical toxicity testing. *Toxicol Sci* 148(1):121–136. <https://doi.org/10.1093/toxsci/kfv171>
- World Health Organization Geneva WHO (2005) Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration-response assessment. IPCS harmonization project document no. 2. <https://www.who.int/publications/i/item/9241546786>. Accessed 28 October 2025
- World Health Organization Geneva WHO (2010) Characterization and application of physiologically based pharmacokinetic models in risk assessment. IPCS harmonization project document no. 9. <https://www.who.int/publications/i/item/9789241500906>. Accessed 28 October 2025
- Yang C, Barlow SM, Muldoon Jacobs KL, Vitcheva V, Boobis AR, Felter SP, Arvidson KB, Keller D, Cronin MTD, Enoch S, Worth A, Hollnagel HM (2017) Thresholds of toxicological concern for cosmetics-related substances: new database, thresholds, and enrichment of chemical space. *Food Chem Toxicol* 109(Pt 1):170–193. <https://doi.org/10.1016/j.fct.2017.08.043>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

SCCS - Scientific Committee on Consumer Safety¹ · Vera Rogiers² · Emilio Benfenati³ · Ulrike Bernauer⁴ · Laurent Bodin⁵ · Patience Browne⁶ · Nicolas Cabaton⁷ · Qasim Chaudhry⁸ · Pieter Jan Coenraads⁹ · Emanuela Corsini¹⁰ · Matthew Dent¹¹ · Corie Ellison¹² · Janine Ezendam¹³ · Eric Gaffet¹⁴ · Corrado Lodovico Galli¹⁵ · Effrosyni Katsanou¹⁶ · Aglaia Koutsodimou¹⁷ · Henriqueta Louro¹⁸ · Abdulkarim Najjar¹⁹ · Alicia Paini²⁰ · Eirini Panteri²¹ · Christophe Rousselle²² · Nikaeta Sadekar²³ · Andreas Schepky²⁴ · Maciej Stępnik²⁵ · Wolfgang Uter²⁶ · Tamara Vanhaecke²⁷ · Natalie von Goetz²⁷ · Susan Wijnhoven²⁸

- ✉ SCCS - Scientific Committee on Consumer Safety
SANTE-SCCS@ec.europa.eu
- ¹ Unit B3: Health Monitoring and Cooperation, Health Networks, Directorate B: Public Health, Cancer and Health Security, European Commission, 2920 Luxembourg, Luxembourg
 - ² Department of In Vitro Toxicology and Dermato-Cosmetology, Vrije Universiteit Brussel, Brussels, Belgium
 - ³ Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy
 - ⁴ German Federal Institute for Risk Assessment (BfR), Berlin, Germany
 - ⁵ Alternative Energies and Atomic Energy Commission (CEA), Saclay, France
 - ⁶ Organisation for Economic Co-operation and Development – OECD, Paris, France
 - ⁷ French National Research Institute for Agriculture, Food and the Environment (INRAE TOXALIM), Toulouse, France
 - ⁸ University of Chester, Chester, UK
 - ⁹ University of Groningen, Groningen, The Netherlands
 - ¹⁰ Università degli Studi di Milano, Milan, Italy
 - ¹¹ Unilever Safety, Environmental and Regulatory Science, Colworth Science Park, Sharnbrook, Bedfordshire MK44, UK
 - ¹² The Procter & Gamble Company, Cincinnati, OH, USA
 - ¹³ National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
 - ¹⁴ Institute Jean Lamour (UMR 7198 CNRS Université de Lorraine), Nancy, France

- ¹⁵ Faculty of Pharmaceutical Sciences, University of Milan, Milan, Italy
- ¹⁶ European Commission, Joint Research Centre (JRC), Ispra, Italy
- ¹⁷ General Chemical State Laboratory, Athens, Greece
- ¹⁸ National Institute of Health Dr. Ricardo Jorge (INSA), Lisbon, Portugal
- ¹⁹ Department Global Toxicology, Beiersdorf AG, Beiersdorfstraße 1-9, 22529 Hamburg, Germany
- ²⁰ European Food Safety Authority (EFSA), Parma, Italy
- ²¹ Faculty of Pharmacy, Laboratory of Pharmaceutical Analysis, National and Kapodistrian University of Athens, Panepistimiopolis - Zografou, 157 71 Athens, Greece
- ²² French Agency for Food, Environmental and Occupational Health & Safety (Anses), European and International Affairs, Maisson-Alfort, France
- ²³ Research Institute for Fragrance Materials - RIFM, Mahwah, NJ, USA
- ²⁴ International Collaboration on Cosmetics Safety – ICCS, Mt. Royal, USA
- ²⁵ Qsar Lab Ltd., Gdańsk, Poland; Department of Toxicology, Faculty of Pharmacy, Medical University of Łódź, Łódź, Poland
- ²⁶ University of Erlangen-Nürnberg, Erlangen and Nuremberg, Germany
- ²⁷ Federal Office of Public Health (FOPH), Bern, Switzerland
- ²⁸ RIVM, Centre for Safety of Substances and Products (VSP), Centre for Substances and Integrated Risk Assessment (SIR), Bilthoven, The Netherlands