Weight of Evidence (WoE) Frameworks: Industry Application

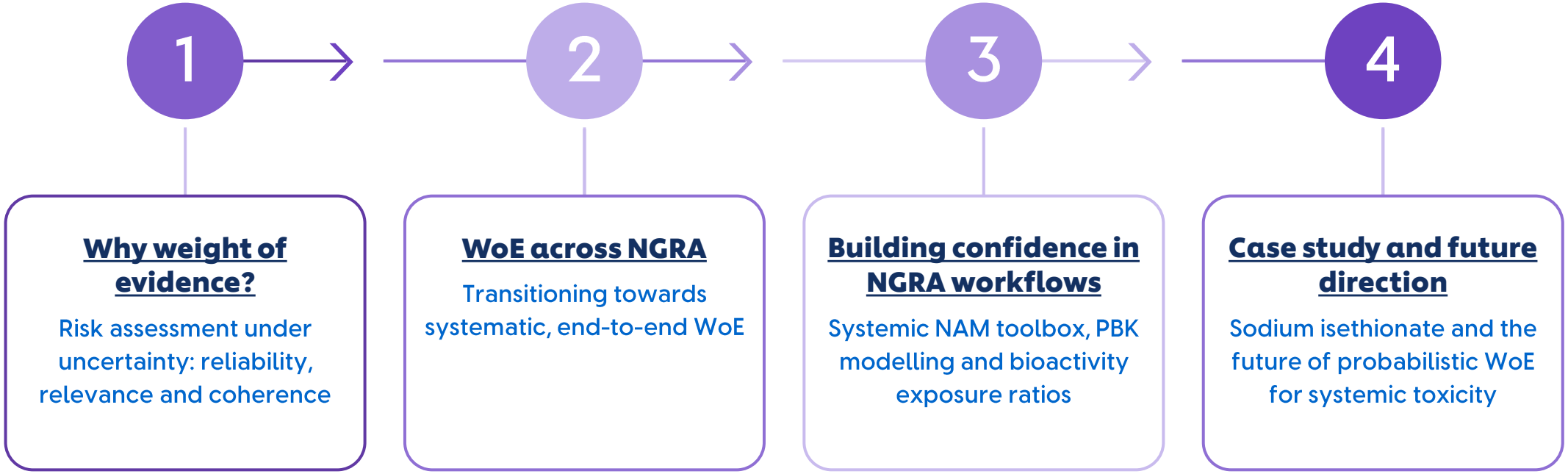
China International Food Safety & Quality Conference (CIFSQ) 2026

Adam Wood: Scientist – Human Safety, Unilever
For any questions contact
adam.wood@unilever.com

SERS
Safety, Environmental & Regulatory Science

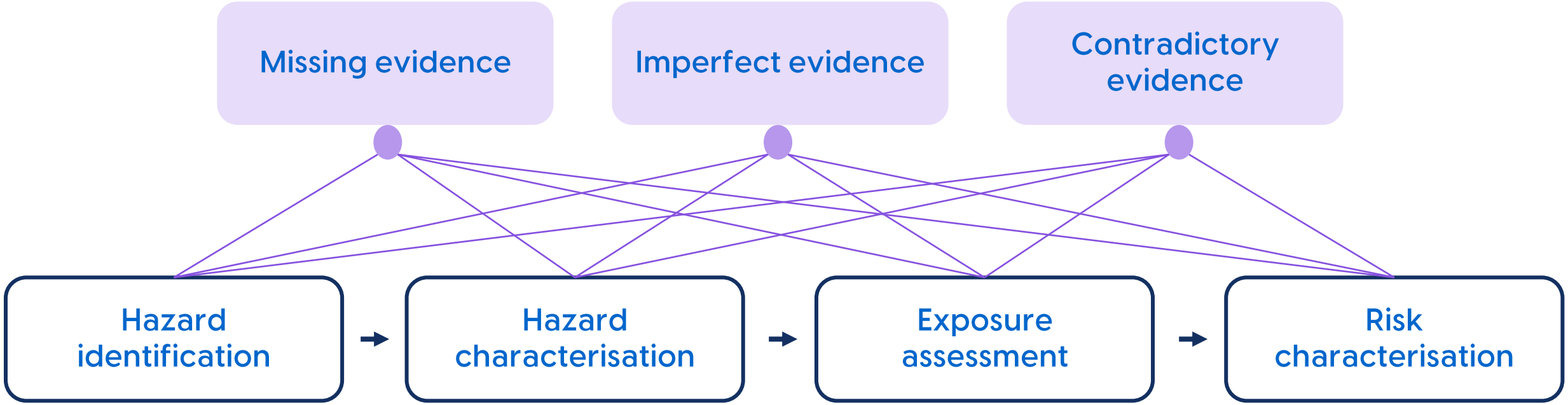
Overview

Unilever



SERS
Safety, Environmental & Regulatory Science

Risk assessment is decision-making under uncertainty



Despite a well-standardised risk assessment framework, each individual step is subject to significant uncertainty. These uncertainties do not exist in isolation and collectively influence the final decision. Risk assessors have the unenviable task of cutting through this complexity

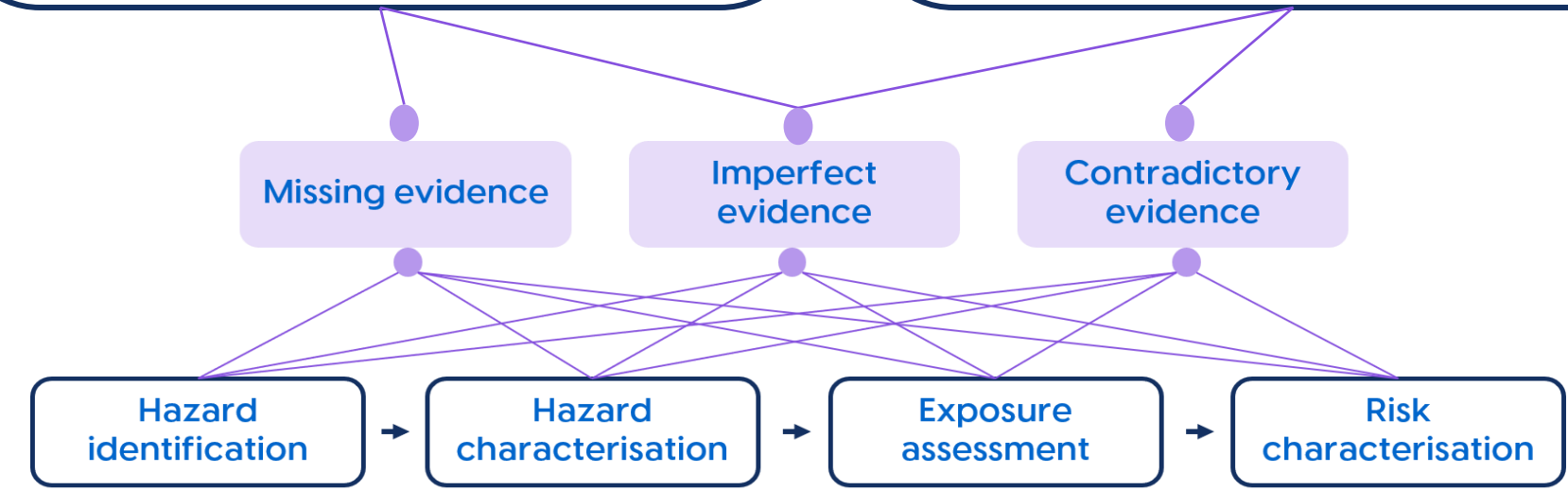
SERS
Safety, Environmental & Regulatory Science

Managing uncertainty in risk assessment

Unilever

Uncertainty factors
Standardised responses to recurring uncertainties, such as:
- Interspecies differences
- Interindividual variability
- Limitations in database completeness

Weight of evidence
Case-specific responses to a particular challenge or question, such as:
- Conflicting results
- Missing or imperfect evidence
- Coherence across evidence streams



SERS
Safety, Environmental & Regulatory Science

What do we mean by weight of evidence?

Unilever

Individual line of evidence

- Reliability
- Relevance

Evidence across strands

- Consistency
- Completeness
- Coherence

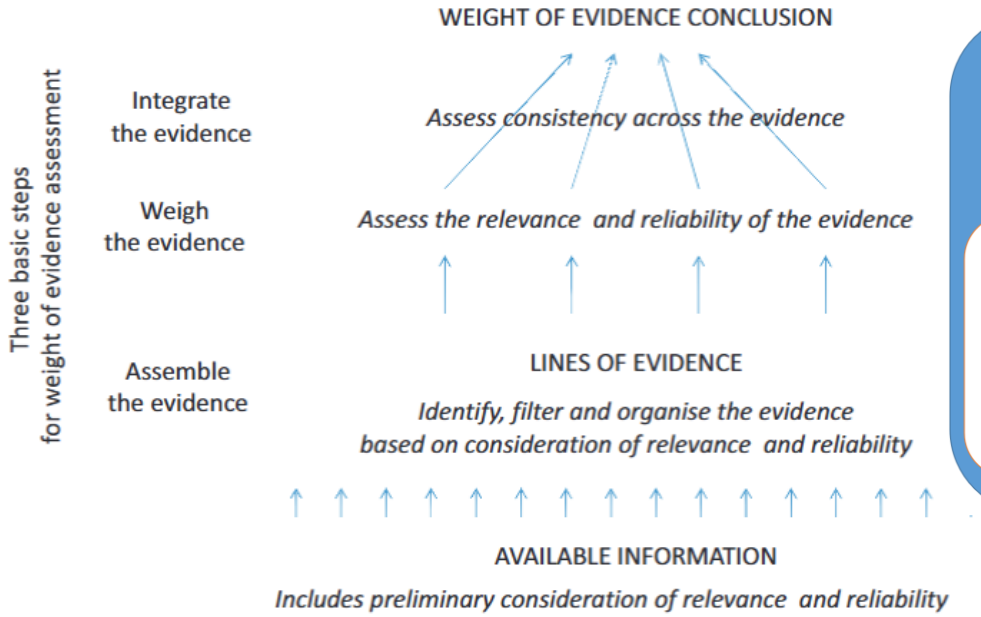
What does the totality of evidence support?

Weight of evidence combines appraisal of individual evidence with integration across evidence streams to determine what the totality of evidence supports.

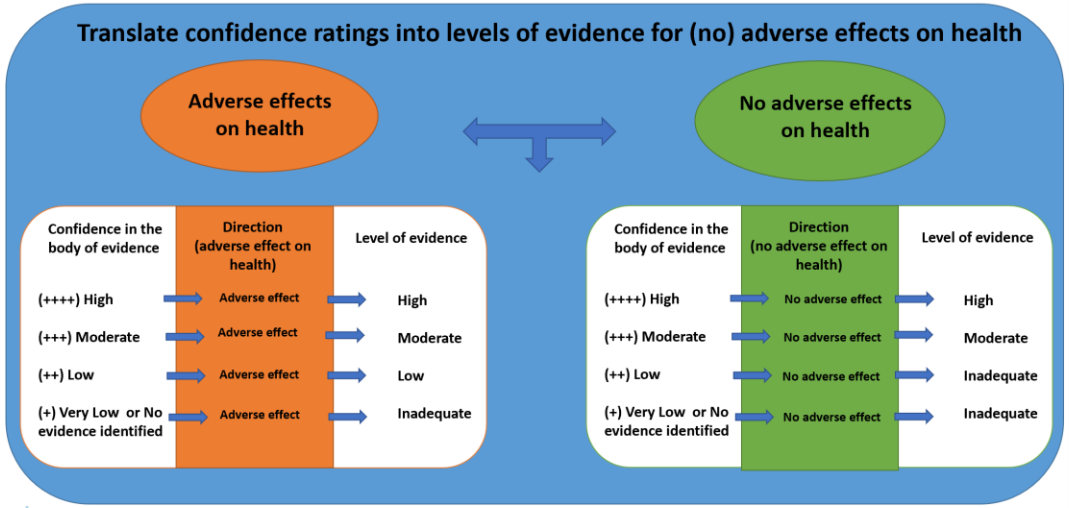
SERS
Safety, Environmental & Regulatory Science

Systematic frameworks improve the transparency of weight-of-evidence

A major application of weight of evidence in the risk assessment community is systematic review, where structured approaches support the consistent appraisal of large and diverse evidence bases. Extensive guidance, case studies and best practice are available.



EFSA, 2017 (Guidance on the use of the weight of evidence approach in scientific assessments)



EFSA, 2022 (Revised Protocol on Hazard Identification and Characterisation of Sweeteners)

SCIENTIFIC OPINION

efsa JOURNAL

Re-evaluation of the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs

EFSA, 2023 (Re-evaluation of the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs)

SERS
Safety, Environmental & Regulatory Science

WoE is not limited to systematic reviews. In next-generation risk assessment, it is a critical end-to-end requirement.

Scientific advances are enabling next-generation risk assessment

Unilever

Recent advances across toxicology, computational science and exposure modelling have substantially increased our ability to demonstrate safety without generating new animal data

Computational approaches

- *In silico* models
- Bioinformatics
- Data science

Human-relevant experimental systems

- *In vitro* and *ex vivo* models
- High-throughput screening
- Omics technologies

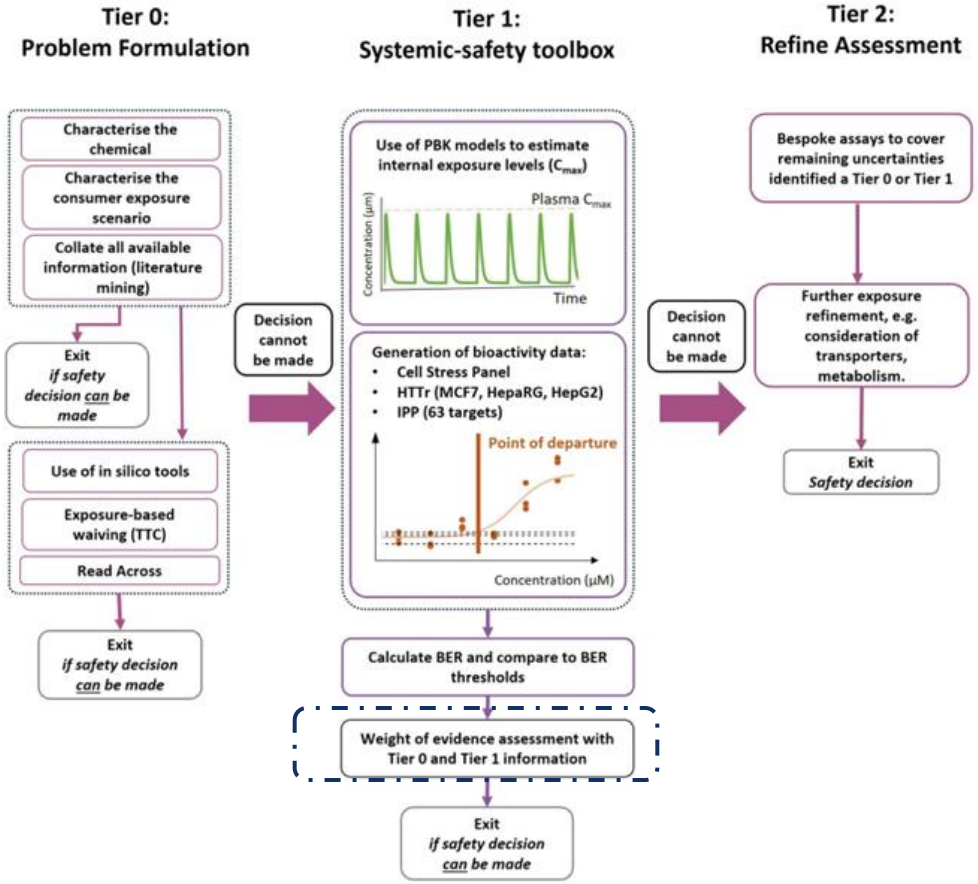
Exposure science

- Physiologically based kinetic modelling

SERS
Safety, Environmental & Regulatory Science

These capabilities provide the building blocks for non-animal next-generation risk assessment (NGRA)

Weight of evidence is an end-to-end requirement in NGRA



OXFORD SOT Society of Toxicology
academic.oup.com/toxsci

Toxicological Sciences, 2025, 204(1), 79–95
https://doi.org/10.1093/toxsci/kfae159
Advance Access Publication Date: December 18, 2024
Research article

Advancing systemic toxicity risk assessment: Evaluation of a NAM-based toolbox approach

Sophie Cable*, Maria Teresa Baltazar, Fazila Bunglawala, Paul L. Carmichael, Leonardo Contreas, Matthew Philip Dent, Jade Houghton, Predrag Kukic, Sophie Malcomber, Beate Nicol, Katarzyna R. Przybylak, Ans Punt, Georgia Reynolds, Joe Reynolds, Sharon Scott, Dawei Tang, Alistair M. Middleton

Safety and Environmental Assurance Centre (SEAC), Unilever, Colworth Science Park, Sharnbrook MK44 1LQ, United Kingdom

*Corresponding author: Safety and Environmental Assurance Centre (SEAC), Unilever, Colworth Science Park, Sharnbrook MK44 1LQ, United Kingdom. Email: sophie.cable@unilever.com

Abstract

For many years, a method that allowed systemic toxicity safety assessments to be conducted without generating new animal test data, seemed out of reach. However, several different research groups and regulatory authorities are beginning to use a variety of in silico, in chemico, and in vitro techniques to inform safety decisions. To manage this transition to animal-free safety assessments responsibly, it is important to ensure that the level of protection offered by a safety assessment based on new approach methodologies (NAMs), is at least as high as that provided by a safety assessment based on traditional animal studies. To this end, we have developed an evaluation strategy to assess both the level of protection and the utility offered by a NAM-based systemic safety "toolbox." The toolbox comprises physiologically based kinetic models to predict internal exposures, and bioactivity NAMs designed to give broad coverage across many different toxicity modes of action. The output of the toolbox is the calculation of a bioactivity:exposure ratio (analogous to a margin of internal exposure), which can be used to inform decision-making. In this work, we have expanded upon an initial pilot study of 10 chemicals with an additional 38 chemicals and 70 consumer exposure scenarios. We found that, for the majority of these (>90%), the NAM-based workflow is protective of human health, enabling us to make animal-free safety decisions for systemic toxicity and preventing unnecessary animal use. We have also identified critical areas for improvement to further increase our confidence in the robustness of the approach.

Keywords: new approach methodologies; risk assessment; systemic toxicity.

A major challenge in the evaluation of nonanimal safety assessments is establishing that they are truly protective of human health. Chemical risk assessment for systemic toxicity has traditionally relied on the use of repeated-dose in vivo tests to derive protective thresholds, mostly based on nonspecific bioactivity observed in animals (Browne et al. 2024). These no-observed (-adverse)-effect-levels (NO(A)ELs) or benchmark dose levels (BMDLs) from chronic and subchronic animal tests are then compared with the estimated exposure in humans. If there is a sufficiently large margin between them, which is determined by using intra- and interspecies uncertainty factors (or other appropriate uncertainty factors arising from consideration of matrices differences, the exposure estimation etc), then the human exposure can be considered low-risk.

The objective of using bioactivity data to derive thresholds that are protective of human health remains the goal of next-generation risk assessment (NGRA) (Dent et al. 2018), i.e. chemical risk assessment that does not involve the generation of new animal data. Where there are data gaps in existing information, or for de novo risk assessments, NGRA assesses bioactivity by estimating points of departure (PoDs) from a range of different in vitro studies and compares them with internal exposure concentrations (e.g. plasma C_{max}) to calculate a bioactivity exposure ratio (BER), analogous to the calculation of a margin of safety/exposure (MoS/E).

The overall approach to such a tiered, nonanimal, exposure-led assessment (Fig. 1) has been described in the literature (Berggren et al. 2017; Thomas et al. 2019), adapted and incorporated into published NGRA case studies (Baltazar et al. 2020; OECD 2021; Ebmeyer et al. 2024; Wood et al. 2024). These case studies demonstrated that NGRA is a flexible approach to integrating multiple lines of evidence to reach a safety decision. However, in order to build confidence in new approach methodologies (NAMs) for systemic toxicity, there is a need to evaluate their robustness and performance in a systematic manner for a wide range of chemicals if they are to be considered for regulatory applications (Dent et al. 2021; Westmoreland et al. 2022). Recent validation frameworks emphasize the need for defining the context of use as a first step in establishing a validation strategy (Van der Zalm et al. 2022; ICCVAM 2024). In our previous work (Middleton et al. 2022), we outlined an approach to evaluating a NAM-based systemic-safety toolbox and workflow comprising bioactivity and exposure modules, which could form a basis for estimating BERs at Tier 1 of an NGRA framework (Fig. 1). Based on the proposed validation frameworks, the "context of use" for this toolbox is defined as an early-tier approach for conducting systemic toxicity risk assessments for consumer goods which aims at being protective of human health, rather than predicting specific hazards, similar to the traditional approach utilising in vivo data (Browne et al. 2024). The

© The Author(s) 2024. Published by Oxford University Press on behalf of the Society of Toxicology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/toxsci/article/204/1/79/7927845 by guest on 23 April 2025

In NGRA, weight of evidence is not a final step. It guides how the assessment is designed, refined and interpreted

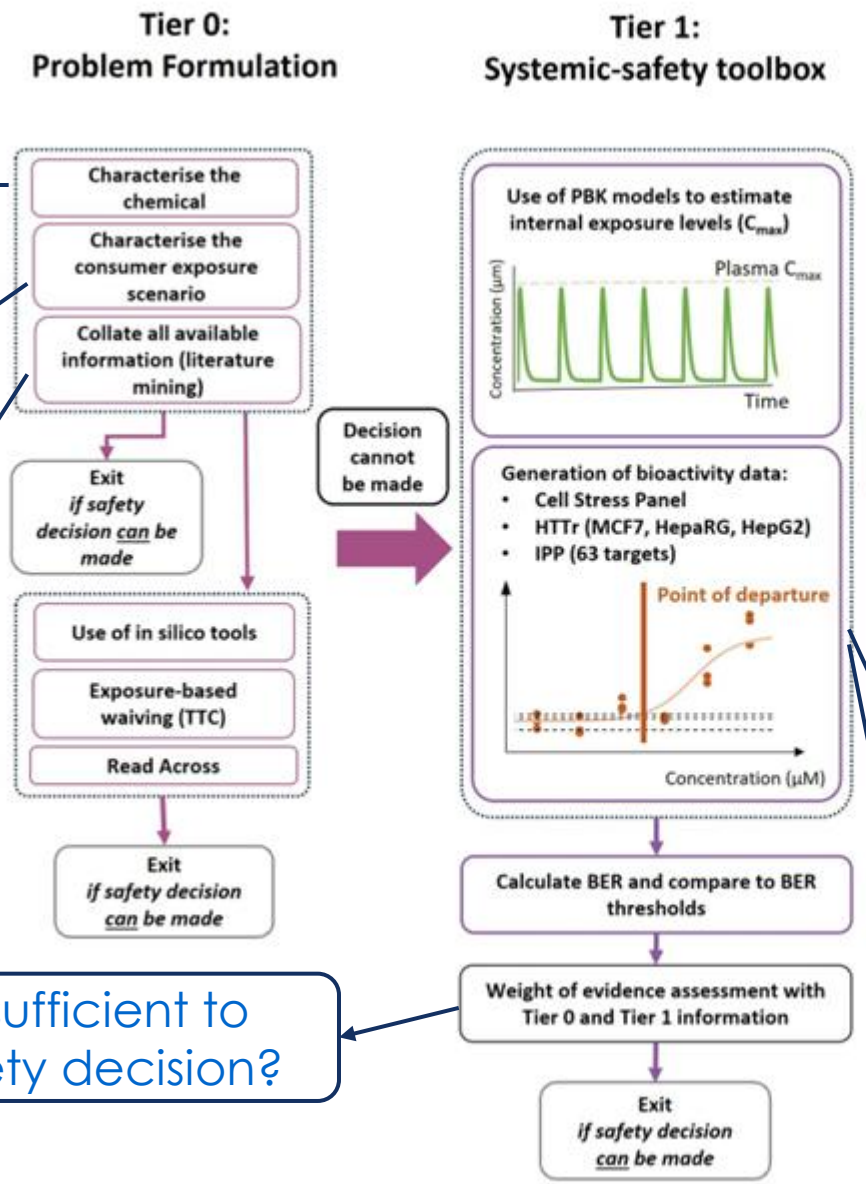
What end-to-end actually means in practice

Unilever

Is the chemical sufficiently characterised to enable hypothesis generation?

Is the exposure scenario sufficiently defined, and are the available data adequate?

What are the existing data telling us, and what question needs answering?



What BER is sufficient to support a safety decision?

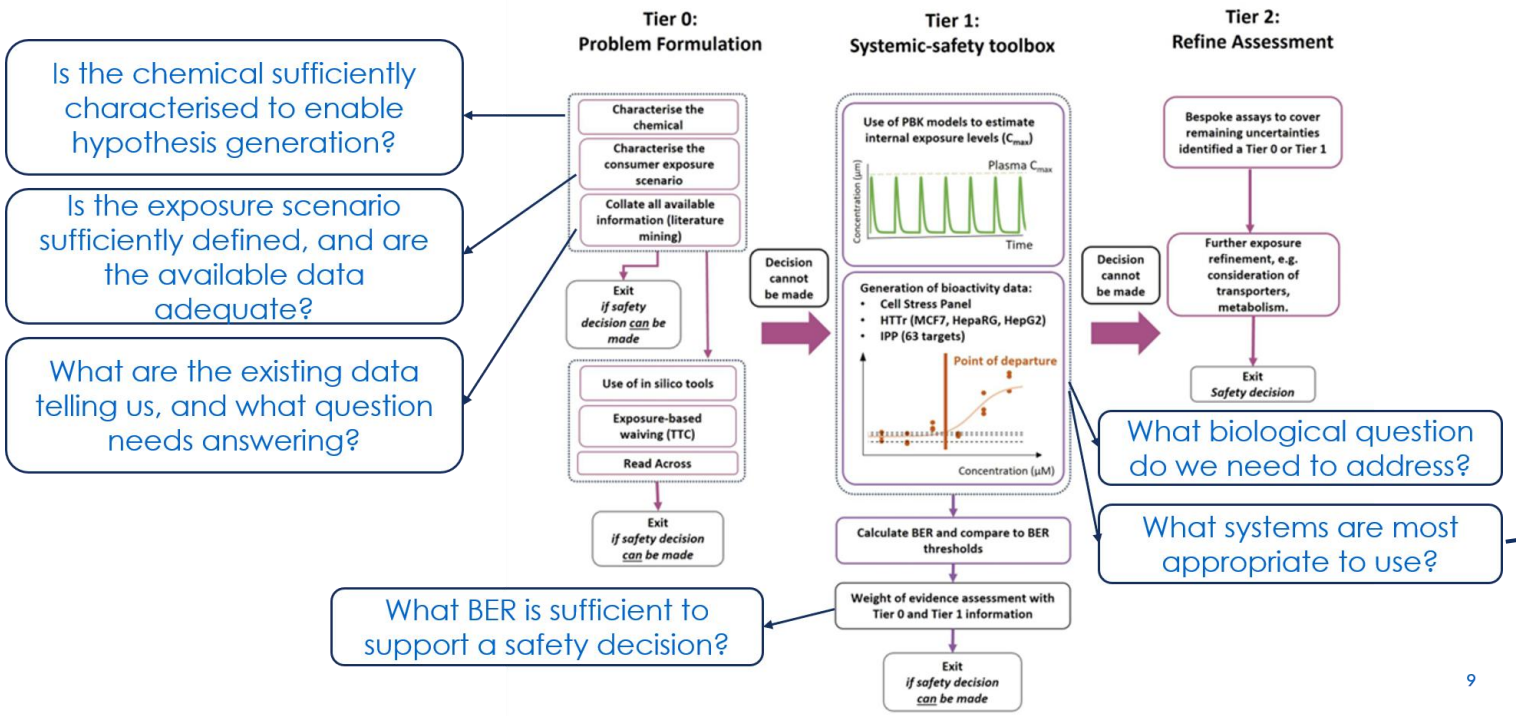
What biological question do we need to address?

What systems are most appropriate to use?

SERS
Safety, Environmental & Regulatory Science

Does limited metabolic competence matter?

Unilever



9

In vitro systems may have limited metabolic competence.

A weight-of-evidence assessment is therefore needed to determine whether metabolism is likely to influence the safety conclusion.

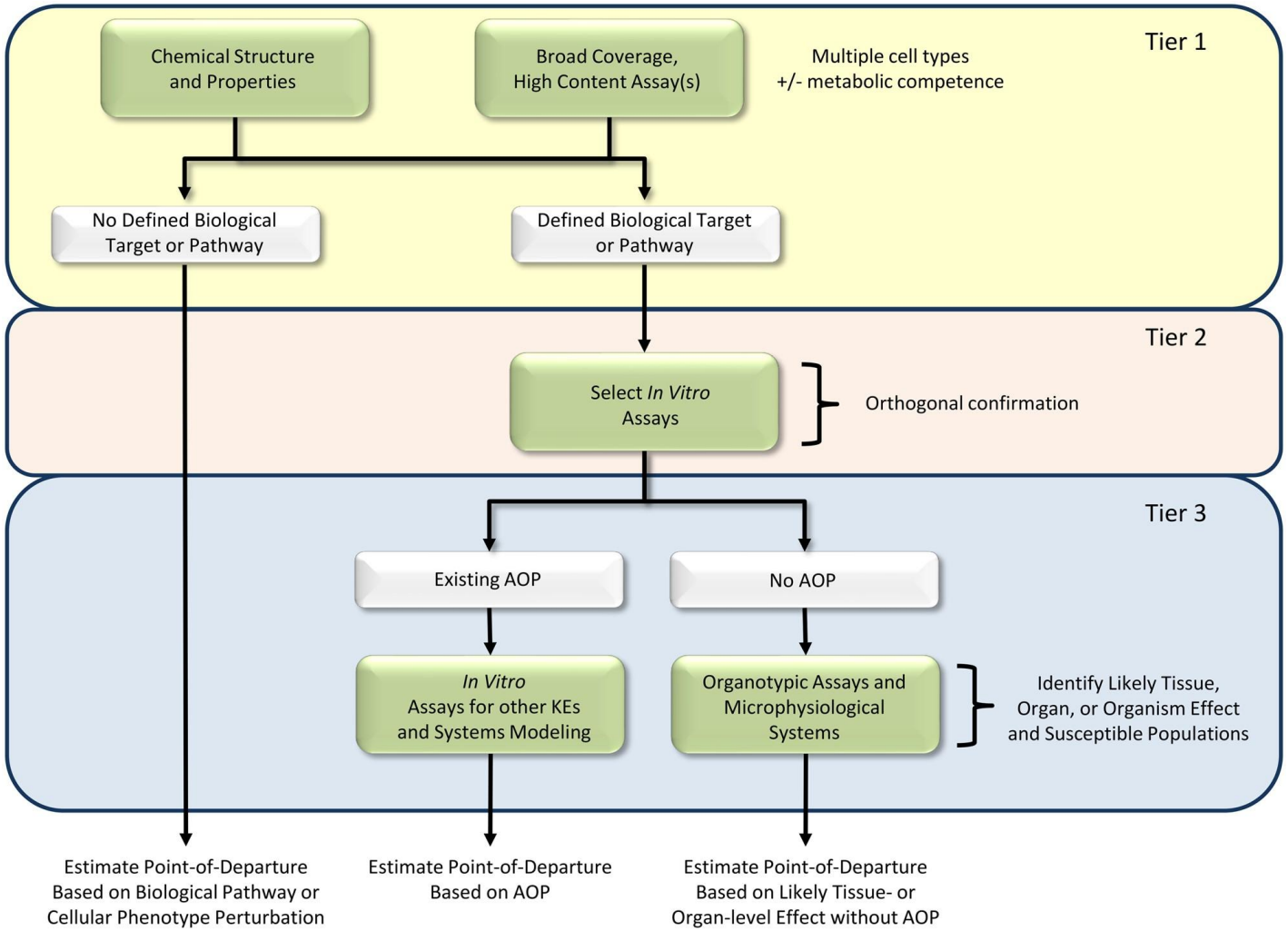
Relevant evidence may include:

- class knowledge
- Expert judgement

A weight-of-evidence assessment is therefore needed to determine whether metabolism is likely to influence the safety conclusion.

SERS
Safety, Environmental & Regulatory Science

Which biological questions should we test?



Toxicol Sci, Volume 169, Issue 2, June 2019, Pages 317–332, https://doi.org/10.1093/toxsci/kfz058

Problem formulation should determine which testing strategy is appropriate.

For many chemicals, the existing evidence may be insufficient to support a sufficiently confident safety conclusion using targeted studies alone.

In these situations, broad-coverage bioactivity NAMs may either support a decision directly or identify uncertainties requiring further investigation.

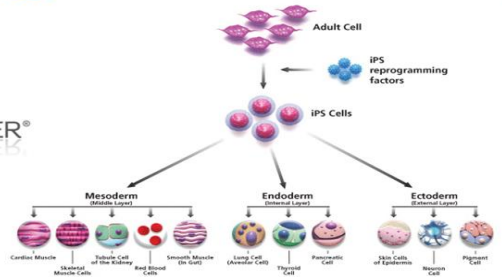
A broad-coverage toolbox for systemic toxicity & DART

Unilever

iPSC based tools

devTOX qP quickPREDICT

REPROTRACKER®

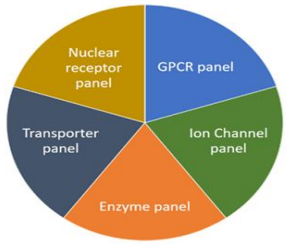


Toxicology in Vitro (2020), **63**, 104746

In vitro Pharmacological Profiling (IPP)

PERSPECTIVES

Reducing safety-related drug attrition: the use of *in vitro* pharmacological profiling

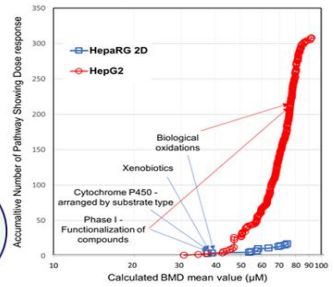


eurofins | Cerep

High-throughput Transcriptomics (HTTr)

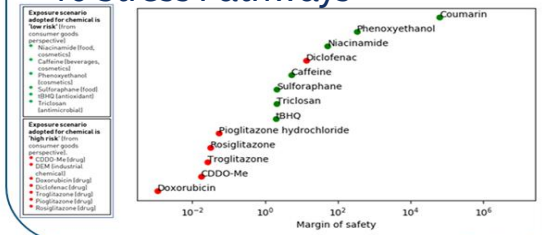
- Use of full human gene panel ~ 21k
- 24 hrs exposure
- 7 concentrations
- 3 cell lines HepG2/ HepaRG/ MCF7
- 3D HepaRG spheroid

BMDExpress 2

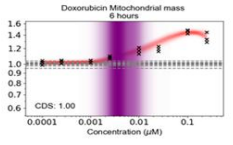


Cell Stress Panel (CSP)

36 Biomarkers, 3 Timepoints, 8 Concentrations – 10 Stress Pathways



cyprotex AN EVOTEC COMPANY



Toxicol Sci (2020), **176**, 11-33

Mueller et al., 2020 | Cable et al., 2024 | Middleton et al., 2022

Complementary targeted and broad-coverage NAMs provide protection across diverse biological mechanisms.

From toolbox to safety decision

Point of Departure (PoD) determination from Bioactivity assays

Unilever

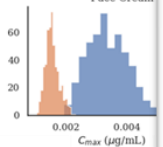
High-Throughput transcriptomics (HTTr)

- TempO-seq technology – full gene panel
- 24hr exposure
- 7 concentrations
- Various cell models (e.g. HepG2, MCF7, HepaRG)
- Dose-response analysis using BMDExpress2 and BIFROST model

Reynolds et al. 2020. Comp Tox 16: 100138
Baltazar et al. 2020. Toxicol Sci 176(1): 236–252

PBK Modelling

Face Cream



Toxicology i

OXFORD SOT Society of Toxicology academic.oup.com/toxsci Tox Spotlight

TOXICOLOGICAL SCIENCES, 189(1), 2022, 124–147

Are Non-animal Systemic Safety Assessments Protective? A Toolbox and Workflow

Alistair M. Middleton, Joe Reynolds, Sophie Cable, Maria Teresa Baltazar, Hequn Li, Samantha Bevan, Paul L. Carmichael, Matthew Philip Dent, Sarah Hatherell, Jade Houghton, Predrag Kukic, Mark Liddell, Sophie Malcomber, Beate Nicol, Benjamin Park, Hiral Patel, Sharon Scott, Chris Sparham, Paul Walker, and Andrew White

ABSTRACT

An important question in toxicological risk assessment is whether non-animal new approach methodologies (NAMs) can be used to make safety decisions that are protective of human health, without being overly conservative. In this work, we propose a core NAM toolbox and workflow for conducting systemic safety assessments for adult consumers. We also present an approach for evaluating how protective and useful the toolbox and workflow are by benchmarking against historical safety decisions. The toolbox includes physiologically based kinetic (PBK) models to estimate systemic C_{ss} levels in humans, and 3 bioactivity platforms, comprising high-throughput transcriptomics, a cell stress panel, and in vitro pharmacological profiling, from which points of departure are estimated. A Bayesian model was developed to quantify the uncertainty in the C_{ss} estimates depending on how the PBK models were parameterized. The feasibility of the evaluation approach was tested using 24 exposure scenarios from 10 chemicals, some of which would be considered high risk from a consumer goods perspective (eg, drugs that are systemically bioactive) and some low risk (eg, existing food or cosmetic ingredients). Using novel protectiveness and utility metrics, it was shown that up to 69% (9/13) of the low risk scenarios could be identified as such using the toolbox, whilst being protective against all (5/5) the high-risk ones. The results demonstrated how robust safety decisions could be made without using animal data. This work will enable a full evaluation to assess how protective and useful the toolbox and workflow are across a broader range of chemical-exposure scenarios.

Key words: Bayesian modelling, new approach methodologies, point of departure, physiologically based pharmacokinetics, probabilistic risk assessment.

The rapid development of new, non-animal approaches for conducting toxicological safety assessments has been driven by several factors. These include ethical considerations, regulatory action (animal test bans for certain types of ingredients), and the need to assure the safety of chemicals using efficient, cost-effective, and robust methods (Dent et al., 2018, 2021; Thomas et al., 2019). Non-animal approaches also have the potential to improve safety assessments by using more human-relevant tools through coverage of key biological pathways or targets. Next-generation risk assessment (NGRA) provides a way to integrate new approach methodology (NAM) data from various sources into the decision-making process, allowing for safety

© The Author(s) 2022. Published by Oxford University Press on behalf of the Society of Toxicology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from academic.oup.com/toxsci/article/189/1/124/6641965 by guest on 09 October 2026

OXFORD SOT Society of Toxicology academic.oup.com/toxsci

Toxicological Sciences, 2025, 204(1), 79–95
https://doi.org/10.1093/toxsci/kfae159
Advance Access Publication Date: December 18, 2024
Research article

Advancing systemic toxicity risk assessment: Evaluation of a NAM-based toolbox approach

Sophie Cable*, Maria Teresa Baltazar, Fazila Bunglawala, Paul L. Carmichael, Leonardo Contreas, Matthew Philip Dent, Jade Houghton, Predrag Kukic, Sophie Malcomber, Beate Nicol, Katarzyna R. Przybylak, Ans Punt, Georgia Reynolds, Joe Reynolds, Sharon Scott, Dawei Tang, Alistair M. Middleton

Safety and Environmental Assurance Centre (SEAC), Unilever, Colworth Science Park, Sharnbrook MK44 1LQ, United Kingdom

*Corresponding author: Safety and Environmental Assurance Centre (SEAC), Unilever, Colworth Science Park, Sharnbrook MK44 1LQ, United Kingdom. Email: sophie.cable@unilever.com

Abstract

For many years, a method that allowed systemic toxicity safety assessments to be conducted without generating new animal test data, seemed out of reach. However, several different research groups and regulatory authorities are beginning to use a variety of in silico, in chemico, and in vitro techniques to inform safety decisions. To manage this transition to animal-free safety assessments responsibly, it is important to ensure that the level of protection offered by a safety assessment based on new approach methodologies (NAMs), is at least as high as that provided by a safety assessment based on traditional animal studies. To this end, we have developed an evaluation strategy to assess both the level of protection and the utility offered by a NAM-based systemic safety "toolbox." The toolbox comprises physiologically based kinetic models to predict internal exposures, and bioactivity NAMs designed to give broad coverage across many different toxicity modes of action. The output of the toolbox is the calculation of a bioactivity:exposure ratio (analogous to a margin of internal exposure), which can be used to inform decision-making. In this work, we have expanded upon an initial pilot study of 10 chemicals with an additional 38 chemicals and 70 consumer exposure scenarios. We found that, for the majority of these (>90%), the NAM-based workflow is protective of human health, enabling us to make animal-free safety decisions for systemic toxicity and preventing unnecessary animal use. We have also identified critical areas for improvement to further increase our confidence in the robustness of the approach.

Keywords: new approach methodologies; risk assessment; systemic toxicity.

A major challenge in the evaluation of nonanimal safety assessments is establishing that they are truly protective of human health. Chemical risk assessment for systemic toxicity has traditionally relied on the use of repeated-dose in vivo tests to derive protective thresholds, mostly based on nonspecific bioactivity observed in animals (Browne et al. 2024). These no-observed (-adverse)-effect-levels (NO(A)ELs) or benchmark dose levels (BMDLs) from chronic and subchronic animal tests are then compared with the estimated exposure in humans. If there is a sufficiently large margin between them, which is determined by using intra- and interspecies uncertainty factors (or other appropriate uncertainty factors arising from consideration of matrices differences, the exposure estimation etc), then the human exposure can be considered low-risk.

The objective of using bioactivity data to derive thresholds that are protective of human health remains the goal of next-generation risk assessment (NGRA) (Dent et al. 2018), i.e. chemical risk assessment that does not involve the generation of new animal data. Where there are data gaps in existing information, or for de novo risk assessments, NGRA assesses bioactivity by estimating points of departure (PoDs) from a range of different in vitro studies and compares them with internal exposure concentrations (e.g. plasma C_{max}) to calculate a bioactivity exposure ratio (BER), analogous to the calculation of a margin of safety/exposure (MoS/E).

The overall approach to such a tiered, nonanimal, exposure-led assessment (Fig. 1) has been described in the literature (Berggren et al. 2017; Thomas et al. 2019), adapted and incorporated into published NGRA case studies (Baltazar et al. 2020; OECD 2021; Ebmeyer et al. 2024; Wood et al. 2024). These case studies demonstrated that NGRA is a flexible approach to integrating multiple lines of evidence to reach a safety decision. However, in order to build confidence in new approach methodologies (NAMs) for systemic toxicity, there is a need to evaluate their robustness and performance in a systematic manner for a wide range of chemicals if they are to be considered for regulatory applications (Dent et al. 2021; Westmoreland et al. 2022). Recent validation frameworks emphasize the need for defining the context of use as a first step in establishing a validation strategy (Van der Zalm et al. 2022; ICCVAM 2024). In our previous work (Middleton et al. 2022), we outlined an approach to evaluating a NAM-based systemic-safety toolbox and workflow comprising bioactivity and exposure modules, which could form a basis for estimating BERs at Tier 1 of an NGRA framework (Fig. 1). Based on the proposed validation frameworks, the "context of use" for this toolbox is defined as an early-tier approach for conducting systemic toxicity risk assessments for consumer goods which aims at being protective of human health rather than predicting specific hazards, similar to the traditional approach utilising in vivo data (Browne et al. 2024). The

© The Author(s) 2024. Published by Oxford University Press on behalf of the Society of Toxicology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/toxsci/article/204/1/79/7927845 by guest on 23 April 2026

o (BER)

og10)

R, the greater the ioactivity will not consumers

SERS
Safety, Environmental & Regulatory Science

A sufficient BER is necessary, but not sufficient

Unilever

Goal: Reach sufficient confidence that the BER supports concluding that the exposure scenario is low risk

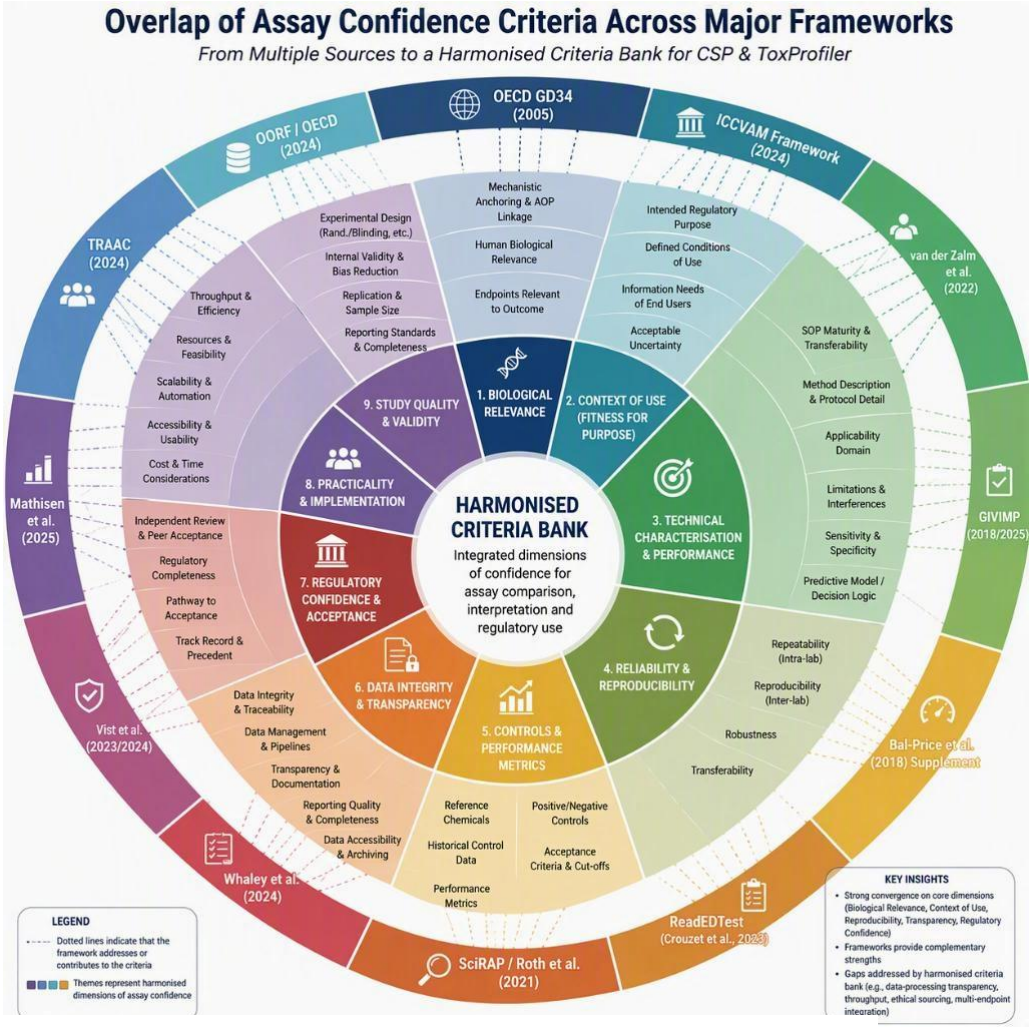
Confidence in initial problem formulation	Confidence in the initial body of evidence retrieved?	Confidence in the integration of the initial body of evidence?
Confidence in exposure assessment	Confidence in the external exposure assessment?	Confidence in the internal exposure assessment?
Confidence in NAM applicability	Confidence in the exposure protocol?	**Confidence in the results**

Building confidence through method-readiness assessment

Unilever

Historical weight-of-evidence frameworks have commonly assessed reliability and relevance by reference to test-guideline status. OECD test-guideline studies are therefore often assigned greater weight. **How should confidence be established for NAMs that lack equivalent formal status?**

Readiness frameworks provide a systematic means of evaluating the strengths and limitations of a NAM. This assessment can inform the confidence placed in the method and the relative weight assigned to its results.

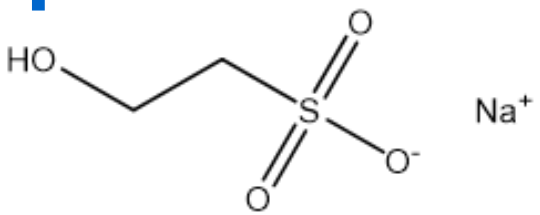


Wood et al., in preparation

SERS
Safety, Environmental & Regulatory Science

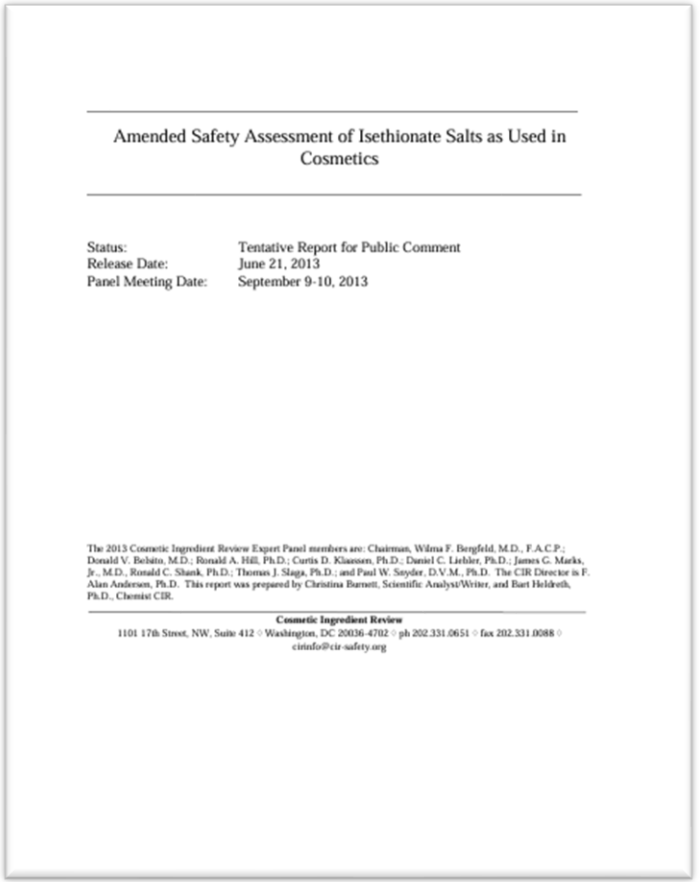
Can NAMs strengthen WoE where regulatory data gaps remain?

Unilever



Sodium 2-hydroxyethanesulfonate (SI) is widely used in the manufacture of alkyl isethionate surfactants.

SI has a long history of use and has been assessed by expert groups, which have concluded that its current use in cosmetic products is of low concern.



Amended Safety Assessment of Isethionate Salts as Used in Cosmetics

Status: Tentative Report for Public Comment
Release Date: June 21, 2013
Panel Meeting Date: September 9-10, 2013

The 2013 Cosmetic Ingredient Review Expert Panel members are: Chairman, Wilma F. Bergfeld, M.D., F.A.C.P.; Donald V. Belsito, M.D.; Ronald A. Hill, Ph.D.; Curtis D. Klaassen, Ph.D.; Daniel C. Liebler, Ph.D.; James G. Marks, Jr., M.D., Ronald C. Shank, Ph.D.; Thomas J. Slaga, Ph.D.; and Paul W. Snyder, D.V.M., Ph.D. The CIR Director is F. Alan Andersen, Ph.D. This report was prepared by Christina Burnett, Scientific Analyst/Writer, and Bart Heldreth, Ph.D., Chemist CIR.

Cosmetic Ingredient Review
1101 17th Street, NW, Suite 412 ○ Washington, DC 20036-4702 ○ ph 202.331.0651 ○ fax 202.331.0088 ○
cirinfo@cir-safety.org

SERS
Safety, Environmental & Regulatory Science

"The Panel concluded that isethionate salts were safe as cosmetic ingredients in the present practices of use and concentration, when formulated to be non-irritating. "

Regulatory data gaps

SI is manufactured and used at several facilities in the EU and is registered under REACH. The applicable tonnage band carries extensive information requirements. SI has a substantial toxicology dataset, but specific regulatory data gaps remain.

Study requirements	< 1 tpa	1–10 tpa Annex VII	10–100 tpa Annex VIII	100–1000 tpa Annex IX	> 1000 tpa Annex X
In vitro skin irritation/corrosion TG 431, TG 439	–	+	+	+	+
In vitro eye irritation TG 467	–	+	+	+	+
In vitro skin sensitisation TG 442E	–	+	+	+	+
In vitro gene mutation in bacteria TG 471	–	+	+	+	+
In vivo acute toxicity: oral TG 401	–	+	+	+	+
In vitro mutagenicity study in mammalian cells or in vitro micronucleus study TG 487, TG 490	–	–	+	+	+
In vitro gene mutation in mammalian cells TG 476	–	–	+	+	+
In vivo pre-natal developmental toxicity in one species TG 414	–	–	–	+	+
In vivo extended one-generation reproductive toxicity (if triggered) TG 443	–	–	–	+	+
In vivo long-term repeated dose toxicity (≥ 12 months) (if triggered) TG 453	–	–	–	–	+
In vivo developmental toxicity in a second species TG 414	–	–	–	–	+
In vivo extended one-generation reproductive toxicity TG 443	–	–	–	–	+
In vivo carcinogenicity if triggered TG 451, TG 453	–	–	–	–	+

Botham et al., *Arch Toxicol* **97**, 3075–3083 (2023).

90-day repeat dose toxicity study: NOAEL 200 mg/kg bw/day

Pre-natal developmental toxicity study (rats): NOAEL >1000 mg/kg bw/day

Can existing evidence, targeted NAMs and exposure-led risk assessment collectively support a sufficiently confident safety conclusion?

(EOGRTS/OECD 443)

2. Pre-natal developmental toxicity study (rabbits)

Existing studies already provide evidence relevant to the data gaps

Unilever

Existing evidence already suggests that the identified data gaps are unlikely to be material to the risk assessment.

SI exposure in an expanded 90-day study was not associated with adverse reproductive-relevant findings, including:

Reproductive organ assessments

- Organ weights
- Macroscopic examination
- Microscopic examination

Functional measurements

- Sperm parameters
- Oestrous cyclicity
- Hormone measurements

SI exposure in a prenatal developmental toxicity study in rats was not associated with maternal or developmental toxicity at the highest dose tested. NOAEL: >1,000 mg/kg bw/day

Initial hypotheses

1. SI is unlikely to produce endocrine-mediated reproductive toxicity.
2. SI is unlikely to produce developmental toxicity at relevant exposures.
3. Targeted NAMs and in silico evidence can test and strengthen these hypotheses.

Mapping occupational exposure across the SI lifecycle

Unilever

- A lifecycle assessment identified relevant uses and associated process categories (PROCs).
- Worker exposures were estimated for these scenarios using appropriate exposure-modelling tools.
- The resulting external exposure estimates were used as inputs to an SI-specific PBK model.

PROC number:	Description:
PROC 1	Chemical production or refinery in closed process without likelihood of exposure or processes with equivalent containment conditions.
PROC 2	Chemical production or refinery in closed continuous process with occasional controlled exposure or processes with equivalent containment conditions
PROC 3	Manufacture or formulation in the chemical industry in closed batch processes with occasional controlled exposure or processes with equivalent containment condition
PROC 4	Chemical production where opportunity for exposure arises
PROC 5	Mixing or blending in batch processes
PROC 7	Industrial spraying
PROC 8a	Transfer of substance or mixture (charging and discharging) at non-dedicated facilities
PROC 8b	Transfer of substance or mixture (charging and discharging) at dedicated facilities
PROC 9	Transfer of substance or mixture into small containers (dedicated filling line, including weighing)
PROC 13	Treatment of articles by dipping and pouring
PROC 14	Tabletting, compression, extrusion, pelletisation, granulation
PROC 15	Use as laboratory reagent
PROC 21	Low energy manipulation and handling of substances bound in/on materials or articles
PROC 28	Manual maintenance (cleaning and repair) of machinery

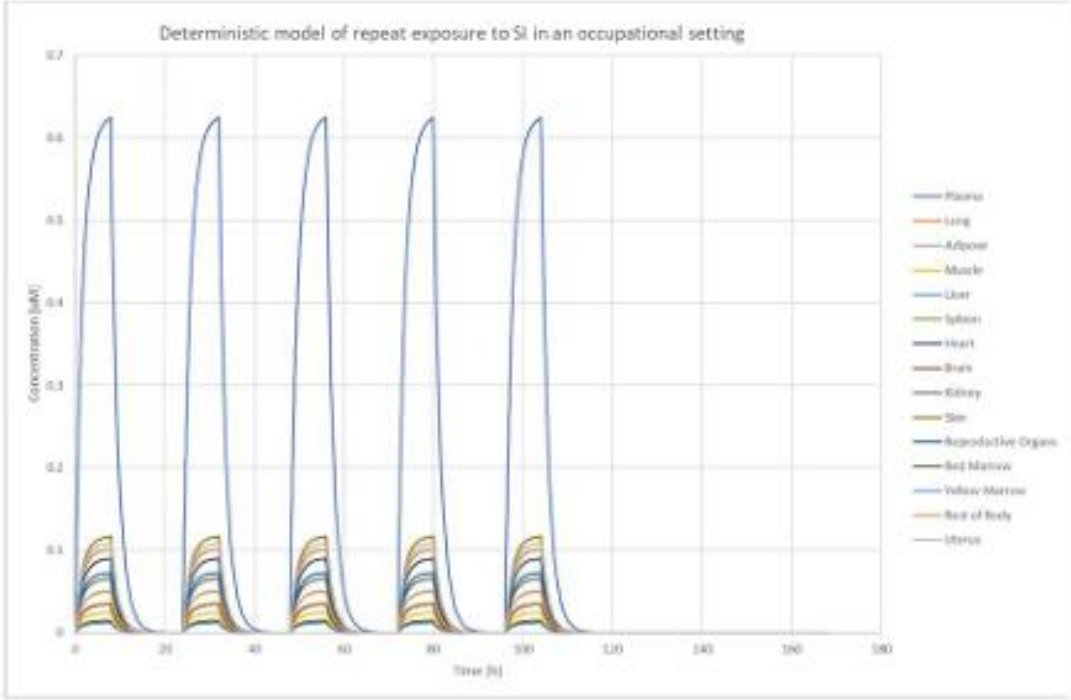
ERM SERS
Safety, Environmental & Regulatory Science

Exposure Scenario	PROC 1	PROC 2	PROC 3	PROC 4	PROC 5	PROC 7	PROC 8a	PROC 8b	PROC 9	PROC 13	PROC 14	PROC 15	PROC 21	PROC 28
Manufacture of substance	☑	☑	☑					☑				☑		☑
Use as Intermediate	☑	☑	☑					☑	☑			☑		☑
Formulation	☑	☑	☑	☑	☑	☑	☑	☑	☑		☑	☑		
Repacking		☑						☑						
Use in Printing inks							☑			☑			☑	
Use as processing aid	☑	☑	☑	☑			☑	☑						
Service Life of fabrics													☑	

PBK modelling translated occupational exposure into internal concentration

- An SI-specific PBK model was developed using experimental ADME data.
- Three simulations represented an individual worker, the general workforce and a pregnant population.
- Probabilistic simulations incorporated uncertainty in model parameters and variability in population physiology.

PBK simulation	C_{max} (µM)	Mean C_{max} (µM)	95th percentile C_{max} (µM)
Single person, deterministic	0.62	-	-
General workforce, probabilistic	-	0.61	0.74
Pregnant population, probabilistic	-	0.58	**0.80**

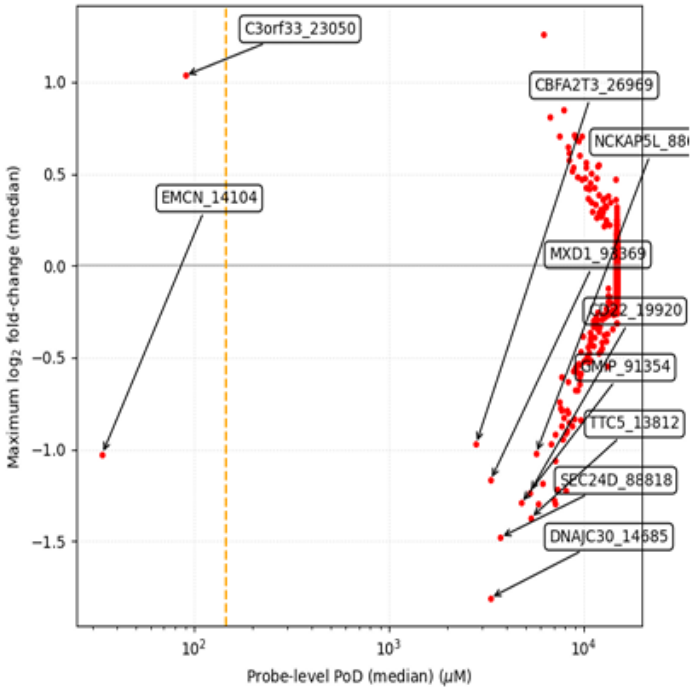


SERS
Safety, Environmental & Regulatory Science

Limited bioactivity across the NAM toolbox

Unilever

- SI showed limited bioactivity across the NAM toolbox.
- The lowest PoD was derived from transcriptomics in MCF-7 cells and was driven by a single probe that was substantially more sensitive than the wider response.
- Dose-confirmation analysis identified deviation from the nominal concentration because of a dosing error.
- The PoD was therefore adjusted using the achieved concentration, increasing confidence in the quantitative *in vitro–in vivo* extrapolation.



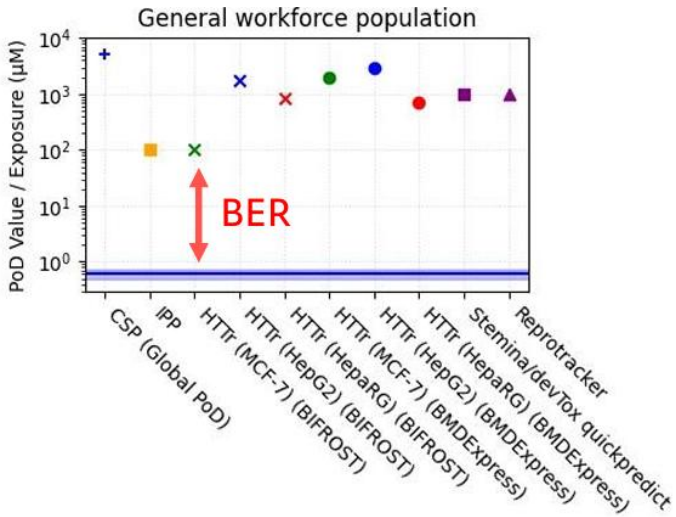
Platform	CSP (Global PoD)	*IPP*	HTTr (MCF-7) (BIFROST)	HTTr (HepG2) (BIFROST)	HTTr (HepaRG) (BIFROST)	HTTr (MCF-7) (BMDExpress)	HTTr (HepG2) (BMDExpress)	HTTr (HepaRG) (BMDExpress)	Stemina/ devTOX quickPredict	Reprotracker
PoD (μM) (Nominal)	7300	>100	150	2500	1200	2860	4210	1040	>1000	>1000
Correction factor based on dose-confirmation study (%)	69.1%	Not determined	69.1%	69.1%	69.1%	69.1%	69.1%	69.1%	None necessary	Not determined
Corrected PoD (μM)	5044	>100	104	1728	829	1976	2909	719	>1000	>1000

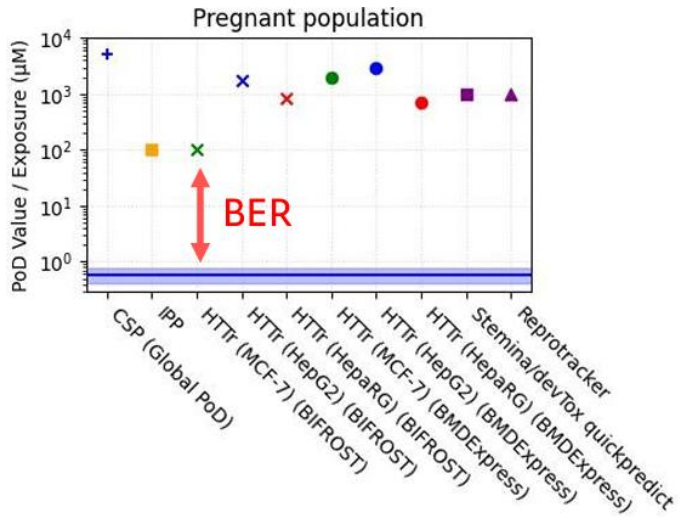
SERS
Safety, Environmental & Regulatory Science

Bioactivity occurred well above predicted internal exposure

Unilever

- The most conservative BER, calculated using the lowest corrected PoD and the pregnant-population 95th percentile Cmax, was 130.
- This exceeded the BER >11 threshold associated with high



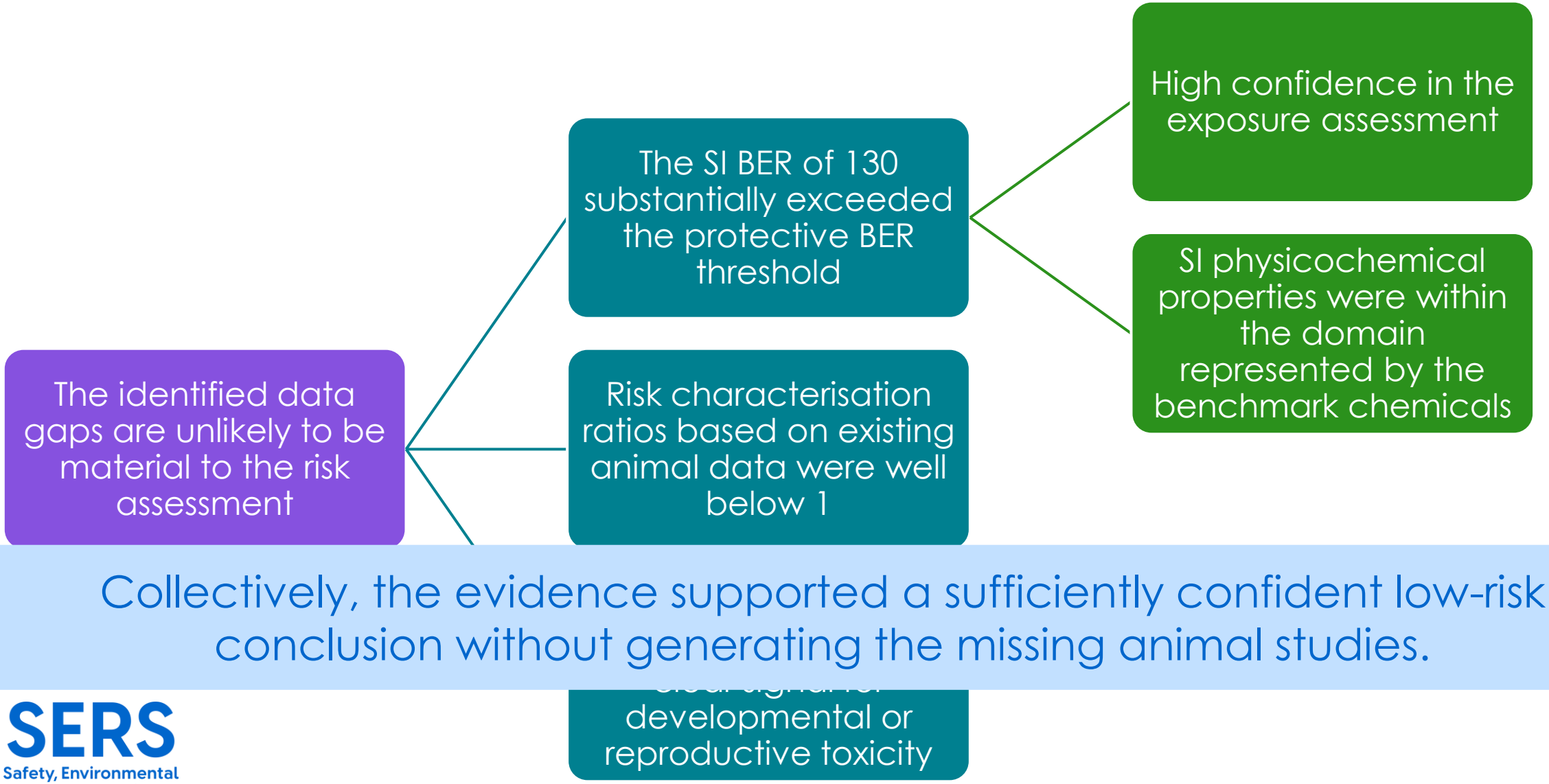


Route	Type of effect	Risk characterisation type	DNEL	PROC 8B Exposure estimate	RCR (ECETOC TRA)	Worst-case BER (ECETOC TRA)

How can we improve the confidence in this decision?

SERS
Safety, Environmental & Regulatory Science

Overall weight of evidence



Collectively, the evidence supported a sufficiently confident low-risk conclusion without generating the missing animal studies.

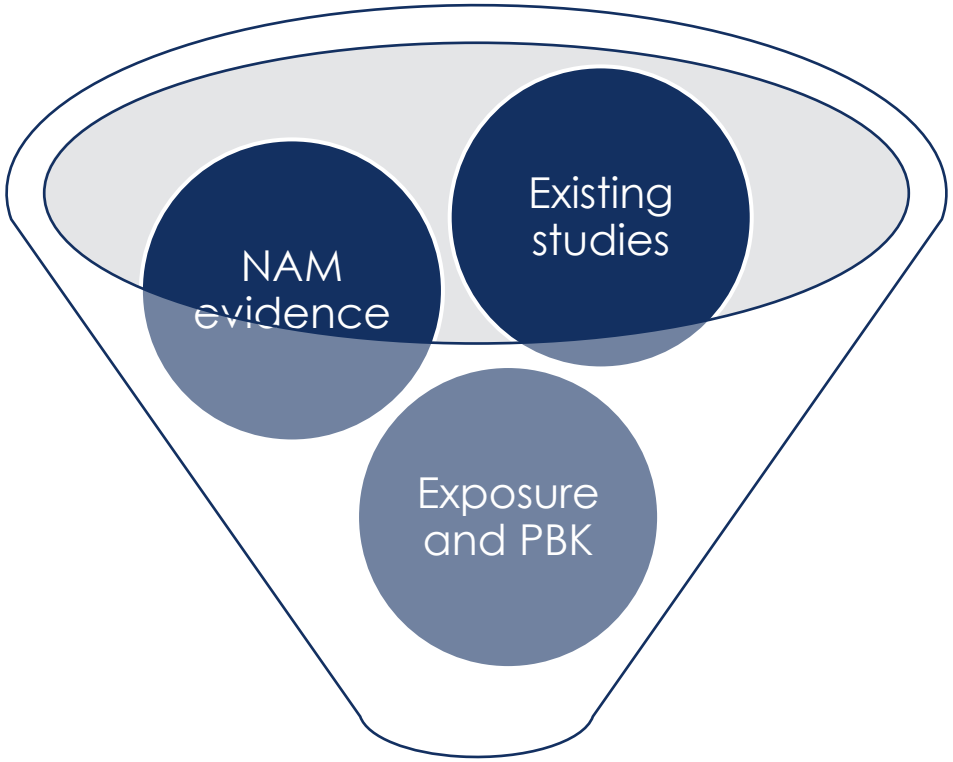
SERS
Safety, Environmental & Regulatory Science

Towards a probabilistic weight-of-evidence framework

Unilever

A probabilistic framework could make the contribution of each evidence stream more explicit, propagate uncertainty through the assessment and quantify how confidence changes as further evidence is added.

Expert judgement would remain necessary, but its assumptions and influence would become more transparent and testable.



Evidence weighting and integration

Updated probability that DART risk is low at the exposure of interest

SERS
Safety, Environmental & Regulatory Science

Wrap-up

Unilever

- Weight of evidence is a **core component** of transparent, evidence-based risk assessment.
- WoE is not confined to integrating evidence at the end of an assessment; **it guides decisions throughout the end-to-end NGRA workflow**.
- Within NGRA, **WoE builds confidence** that the available evidence is sufficiently reliable, relevant and coherent to support a safety decision.
- Probabilistic frameworks offer a route towards more explicit, transparent and reproducible evidence integration, while retaining expert judgement.

Leading safety and environmental sustainability sciences
The scientists behind our safe and sustainable products.

Safe and sustainable by design
How we build safety and sustainability into every product innovation.

Keeping people and the environment safe
The science-based approaches we use to keep our consumers, workers and the environment safe.

Reducing our environmental impact
How we harness the latest science to minimise our environmental footprint.

Learn more about Unilever Safety and Environmental Science
https://www.unilever.com/sustainability/safety-and-environment/

SERS
Safety, Environmental & Regulatory Science