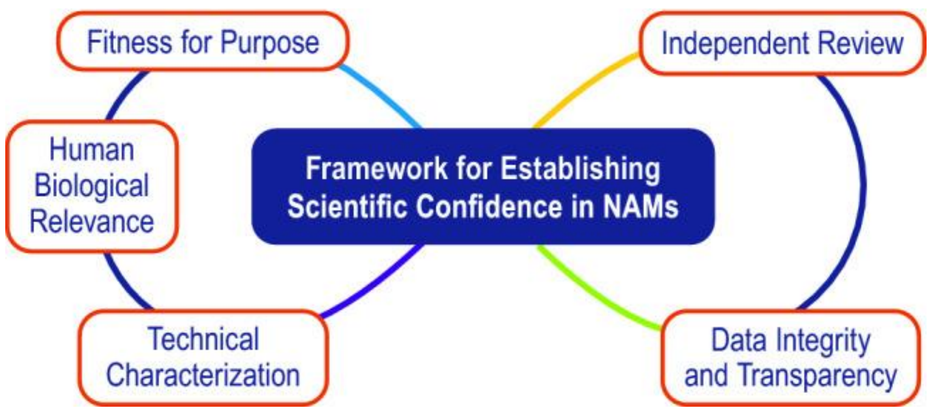
Establishing scientific confidence in a NAM toolbox and workflow using a flexible framework

SEAC | **Unilever**

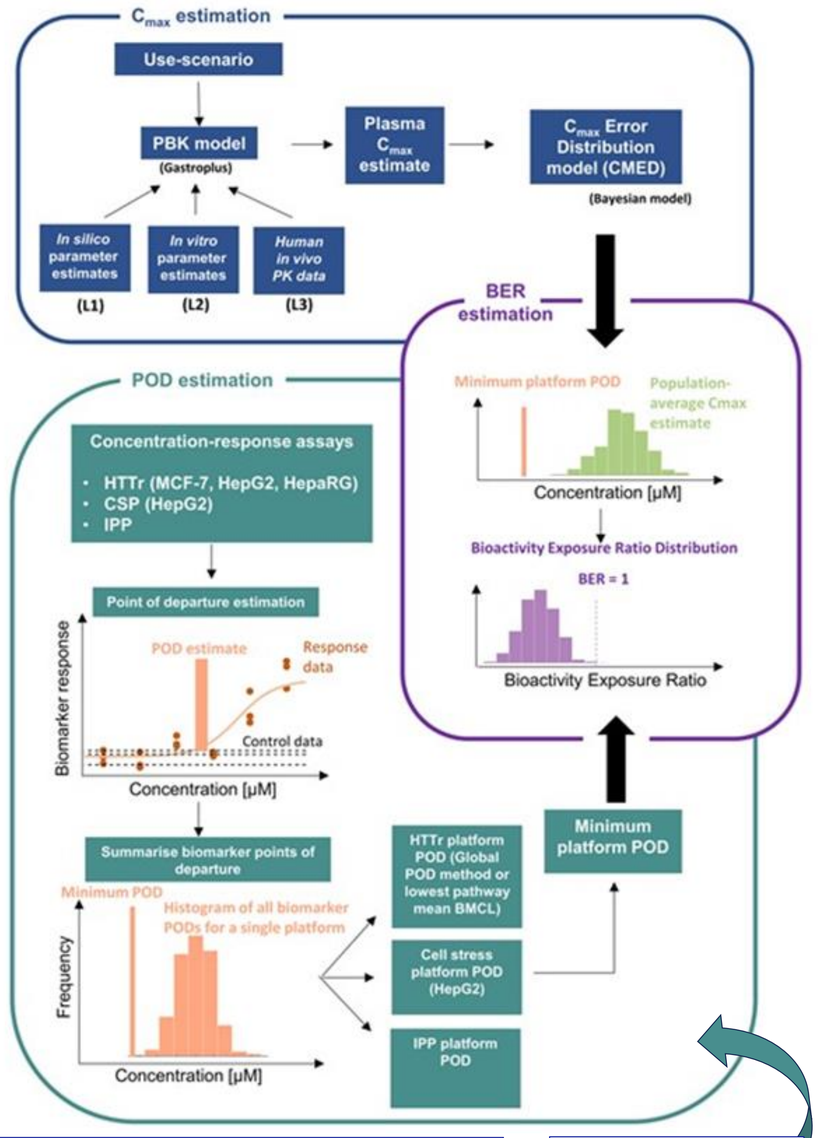
Reynolds, G., Cable, S., Carmichael, P., Dent, M., Kukic, P., Malcomber, S., Middleton, A., Przybylak, K., Punt, A., Reynolds, J., Scott, S., White, A., Baltazar, M.

Introduction

In recent years significant progress has been made in the development, evaluation, and application of new approach methods (NAMS) for next generation risk assessment (NGRA) of systemic safety, which is increasing confidence in their use for making robust safety decisions. However, it is important to go beyond this and evidence areas such as technical characterization of decision frameworks and their component NAMs, to establish scientific confidence for regulatory purposes. In a paper by Van der Zalm et al., (2022) a framework for establishing confidence in NAMs was proposed, comprising five elements (fitness for purpose, human biological relevance, technical characterization, data integrity and transparency, and independent review). This flexible approach was applied to the components of the systemic toolbox and workflow described in Middleton *et al.* (2022). The toolbox - intended to be used as a tier one approach within an integrated approach to testing and assessment (IATA) for cosmetic safety assessments - includes physiologically based kinetic (PBK) models to estimate human plasma C_{max}, and 3 bioactivity platforms, comprising high-throughput transcriptomics (HTTr), a cell stress panel (CSP), and *in vitro* pharmacological profiling (IPP), from which points of departure (PoD) are estimated and a bioactivity exposure ratio (BER) defined. This poster demonstrates it is possible to expand upon a flexible framework to establish confidence in a multi-NAM workflow, with evidence for a fit for purpose, robust, relevant and reliable approach, going beyond a defined endpoint and instead focussing on overall protection of human health. The framework for scientific confidence also provides a scaffold to easily identify where evidence may be lacking, so that strategies can be implemented to fully meet these principles.



←*Figure 1: Framework for establishing scientific confidence, published in Van der Zalm et al., (2022). Reproduced with permission.*



→*Figure 2: Schematic of the systemic safety toolbox and associated workflow, which comprises 3 modules: one to estimate the exposure using physiologically based kinetic (PBK) models, another to estimate the point of departure (POD) based on the cell stress panel (CSP), high throughput transcriptomics (HTTr), and in vitro pharmacological profiling (IPP) bioactivity data. The workflow involves combining the outputs from these 2 modules into the third module to estimate the bioactivity exposure ratio (BER). Published in Middleton et al., (2022).*

Results

Context of Use

What is the context of use?

'Tier 1' PoD and BER approach to systemic risk assessment, for use within a tiered and iterative framework for an integrated approach to testing and assessment (IATA). Developed for safety assessment of consumer goods; however, it should not necessarily be limited to this industry.

Examples of application of the approach within this context include Baltazar et al., 2020, and OECD, 2021.

What information does the NAM provide?

Exposure: Human plasma C_{max}
Bioactivity PoDs:
- high-throughput transcriptomics (HTTr)
- cell stress panel (CSP)
- *in vitro* pharmacological profiling (IPP)

BER: ratio of global PoD (lowest PoD across the combination of NAMs) to C_{max}
Decision Model: *Low Risk* **=** BER > threshold *Uncertain risk*: BER < threshold

How does the information relate to the regulatory need?

This toolbox is **not** a one-to-one replacement for systemic endpoints as defined by the current animal-based hazard and/or risk standards defined in regulatory statutes.

Use of the toolbox in the context of a tiered framework is intended to negate the need for *in vivo* repeated dose toxicity studies by a protection-not-prediction approach to risk assessment. Much like the *in vivo* studies, these NAMs are not intended to establish mechanism of toxicity. Instead of using a threshold for a general apical effect e.g. body weight change, the toolbox and workflow uses a measure of bioactivity as the PoD.

How is the data incorporated?

Individual assays are incorporated within a toolbox and workflow as per Fig 2. This toolbox is intended to be used as part of a tiered and iterative risk assessment framework.

Technical Characterisation

Relevant Method

Is the approach human relevant?

Consideration has been given to the human relevance of the approach via multiple sources. For example, **human cell lines** are utilised throughout the toolbox assays. 44 of the IPP panel targets have been associated with *in vivo* adverse drug reactions [Bowes et al., 2012]. I*n vitro* hepatic intrinsic clearance data are generated in **primary human hepatocytes** for PBK modelling due to superior performance vs other *in vitro* methods (Chiba et al., 2009). PBK models are calibrated using human clinical data where possible (Moxon *et al.*, 2020). **'True Dose'** i.e. implications of volatility, stability, hydrophobicity, binding to plastic and serum and solubility are considered to build confidence in the biologically effective dose in *in vitro* test systems; for the majority of chemicals the use of nominal concentrations would result in conservative decision making (Nicol et al., 2024).

Applicability Domain

Chemical space: logP for 38 test chemicals (Cable et al., 2024, submitted) ranged from 2.32-5.99, molecular weights ranged from 100.12-780.9 g/mol. Chemical space analysis by molecular descriptors and structural fingerprints demonstrated coverage of chemical space, however, there are structural regions not represented, such as structures with long (C8 and above) alkyl linear moieties.

Biological coverage: by having a multi-NAM approach that covers cell stress pathways (CSP), key pharmacological targets (IPP) and broad coverage through HTTr using multiple cell lines, biological coverage is increased. In addition, the evaluation chemical test set contains chemicals which are not associated with any adverse effects at relevant exposures, some with local effects, some associated with non-specific and some with specific toxicities. Where evidence was found for effects in humans, effects included hepatoxicity, immune modulation, blood-based disorders, nervous system disruption, neurological effects, cardiac effects, nephrotoxicity, gastrointestinal issues, and an example of pulmonary fibrosis.

Exposure routes: modelled scenarios included; 2 exposures via the inhalation route, 7 via intravenous administration, 11 via dermal application and 32 via oral administration.

Equivalence of Decision

Depending on the PBK level used to parameterise the model, the toolbox is protective of 93% (*in silico*), 93% (*in vitro*) and 98% (clinical) of high-risk benchmark chemical exposures. The utility of the approach, i.e. correctly identifying low-risk benchmark chemical exposures, was 8%, 24% and 0% respectively. Protectiveness and utility of the traditional approach was calculated to be 97% and 42% when using the lowest in vivo NOEL/NOAEL and a Margin of Safety of 100.

For the same chemicals, the performance of the NAM-based toolbox was roughly equivalent (96% protectiveness and 32% utility)".

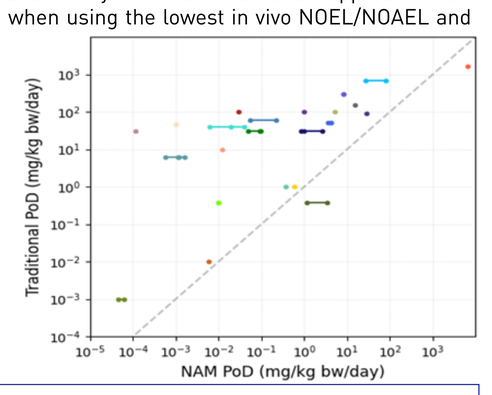


Figure 3. Comparison of equivalent external dose NAM PoDs transformed from μM to mg/kg bw/day vs. minimum traditional PoDs (mg/kg bw/day). Pearson correlation 0.57. For nearly all chemicals the NAM PoD was more conservative

Robust Method

Evaluate assay protocol

HTTr and IPP methodologies and evaluations have been published (Bowes et al., 2012; Reardon et al., 2023; Harrill et al., 2024). Development and evaluation of the CSP was documented in Hatherell et al., 2020. Middleton et al., 2022 details experimental design finalisation for subsequent toolbox evaluation submitted in Cable et al..

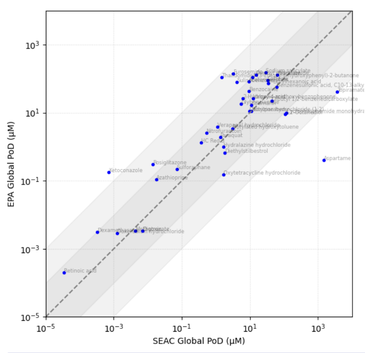
Evaluate Computational Methods

BIFROST, used to calculate the 'global POD' for HTTr and CSP, and exposure estimates i.e. PBK C_{max} in the CMED model has been published (and code shared) in Middleton et al., 2022, Reynolds et al. 2020, Hatherell et al., 2020. **Model Reviews**: A 'Bayesian Workflow' (Gelman et al., 2020) summarises the steps for evaluating and validating this type of model. Aspects covered in this framework include; posterior predictive checks, cross validation, influence of individual data points, influence of prior, prediction and convergence diagnostics. **Code Reviews & Testing**: Code development, review and testing should be iterative. Unit and integration tests must be implemented by the developers to ensure that each function, and the code as-a-whole, does what is expected. Formal end user acceptance testing (UAT) should be conducted. Finalisation and publication of these aspects is forthcoming.

Reliable Method

Method Variability and Transferability

A preliminary dataset for chemical reproducibility of PoDs in reduced cell stress panel assay, showed a range up to 3-fold for the majority of chemicals and up to 9.4-fold for caffeine and niacinamide – this wider variability could possibly be due to the solubility of the top stock concentrations with the differences in preparation. Correlation of PoDs between laboratories was high (0.97). It was determined that the PoDs obtained from the Unilever data are as consistent with PoDs derived from the CRO (Cyprotex) data, as Cyprotex PoDs are with themselves.



→ *Figure 4. A correlation of 0.82 was found when similar BIFROST models were used to analyse HTTr HepG2 datasets generated at EPA versus Unilever CROs (BioClavis/Cyprotex). Aspartame is an outlier, however it's HTTr signal is distinctly the same as retinoic acid, which was adjacent on the dosing plate, and therefore suspected to be contaminated with high confidence.*

Data Integrity and Transparency

Is an appropriate quality system applied to the approach?

Each of the labs responsible for the bioactivity assays within the toolbox uses their *own quality systems*. These quality systems all document in the following areas; facility management responsibilities, personnel and training, auditing, facilities, equipment and systems, materials and reagents, test systems, test and reference items, SOPs, protocols, data, reporting, archive. Experimental technical documentation is provided for the cell stress panel and HTTr cell lysate generation.

Application of the approach has undergone some independent review via an OECD Series on Testing & Assessment case study on Phenoxyethanol and ongoing discussions with SCCS on a case study for BP4. Individual components of the toolbox have not yet been subject to independent technical review e.g. for BIFROST model and code review.

Transparency

Publications which support all areas of this approach are referenced throughout this poster and we continue to share evaluation results, code, methods and examples of application of the approach through peer reviewed publications.

Independent Review

Remaining gaps and next steps

Whilst initial evaluation of the toolbox and workflow has been completed, all aspects of a flexible, NAM relevant, validation framework are yet to be completed. However, this review of the current state of play demonstrates that aspects of all elements can, and in many cases have, been fulfilled.

Some elements still to be exemplified are;

- Publication of BIFROST model and code review.
- Individual components of the toolbox have not yet been subject to a full independent review
- Further reproducibility and transferability studies beyond proof of principle studies

In addition, ongoing technical improvements to the toolbox will be pursued e.g. to improve the utility of the toolbox including optimisation and standardisation of the CSP and expansion of the dataset to more chemical classes.

References

Baltazar *et al.* 2020. A Next-Generation Risk Assessment Case Study for Coumarin in Cosmetic Products. Toxicological Sciences. 178(1):236-252

Cable *et al.*, 2024. Submitted. Advancing systemic toxicity risk assessment: Evaluation of a NAM-based toolbox approach.

Gelman *et al.*, 2020. Bayesian workflow. arXiv preprint arXiv:2011.01808.

Hatherell *et al.*, 2020. Identifying and Characterizing Stress Pathways of Concern for Consumer Safety in Next-Generation Risk Assessment. Toxicological Sciences. 176(1): 11-33

OECD, 2021. Case Study on the Use of the use of Integrated Approaches for Testing and Assessment for the Systemic Toxicity of Phenoxyethanol when included at 1% in a body lotion, Series on Testing & Assessment No. 349. ENV/CBC/MONO(2021)35, OECD, Paris

OECD, 2018. Guidance Document on Good In Vitro Method Practices (GIVIMP).

Middleton *et al.*, 2022. Are non-animal systemic safety assessments protective? A toolbox and workflow. Toxicological Sciences, 189(1), 124-147.

Moxon *et al.*, 2020. Application of physiologically based kinetic (PBK) modelling in the next generation risk assessment of dermally applied consumer products. Toxicology in Vitro, 63, 104746.

Van Der Zalm *et al.*, 2022. A framework for establishing scientific confidence in new approach methodologies. Archives of Toxicology, 96(11), 2865-2879.