

**Next Generation Risk Assessment (NGRA)
using New Approach Methods (NAMs)
to Evaluate Systemic Safety for Consumers
using Benzophenone-4 as a UV-filter in a
Sunscreen Product**

**Maria Baltazar,
Unilever**

01/02/2023

We personally care



Cosmetics Europe
the personal care association

Benzophenone-4 (BP-4) case study: Objectives & Approach

In 2019, the European Commission defined a list of 28 cosmetic ingredients with potential endocrine activity

BP-4 is one of the 28 chemicals for which the call for data took place.

Objective of the case study:

- **To assess whether a tiered NGRA approach is sufficiently protective and also useful to answer a real-life question**

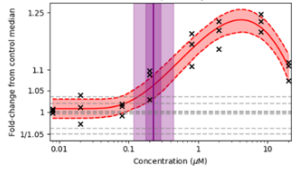
Is Benzophenone-4 safe in a sunscreen product at the maximum approved level of 5%?

Benzophenone-4 (BP-4) case study: rules & assumptions

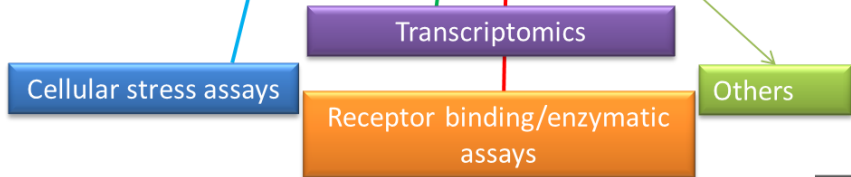
- Focus on systemic toxicity
- For the purposes of this exercise, it has been assumed that no in vivo animal data exist on the ingredient;
- Stand-alone illustration of how to assess systemic toxicity effects (not including genetic toxicity) using NAMs and does not rely on being used in conjunction with any previous dossiers

Approach to this Next Generation Risk Assessment – Protection of human health

Point of departure (POD) derived from concentration-response data



Systemic toolbox of assays (NAMs) which cover a broad biological space – measurements of bioactivity



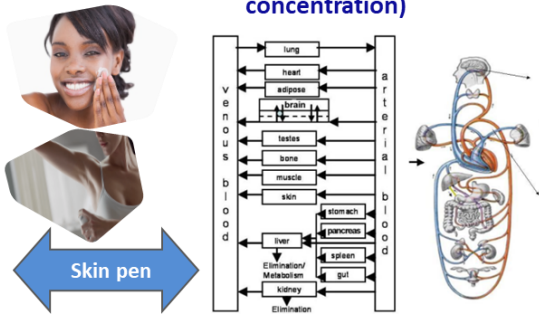
If there is no bioactivity observed at consumer-relevant concentrations, there can be no adverse health effects.

Calculation of Bioactivity exposure ratio (BER)

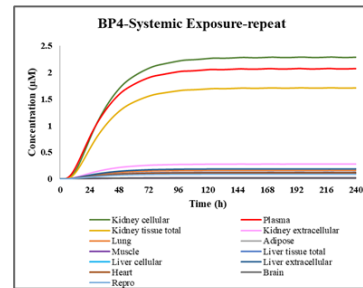
The BER is defined as the ratio between the POD and the relevant exposure metric

If there is bioactivity observed at consumer-relevant concentrations -> is it adverse?

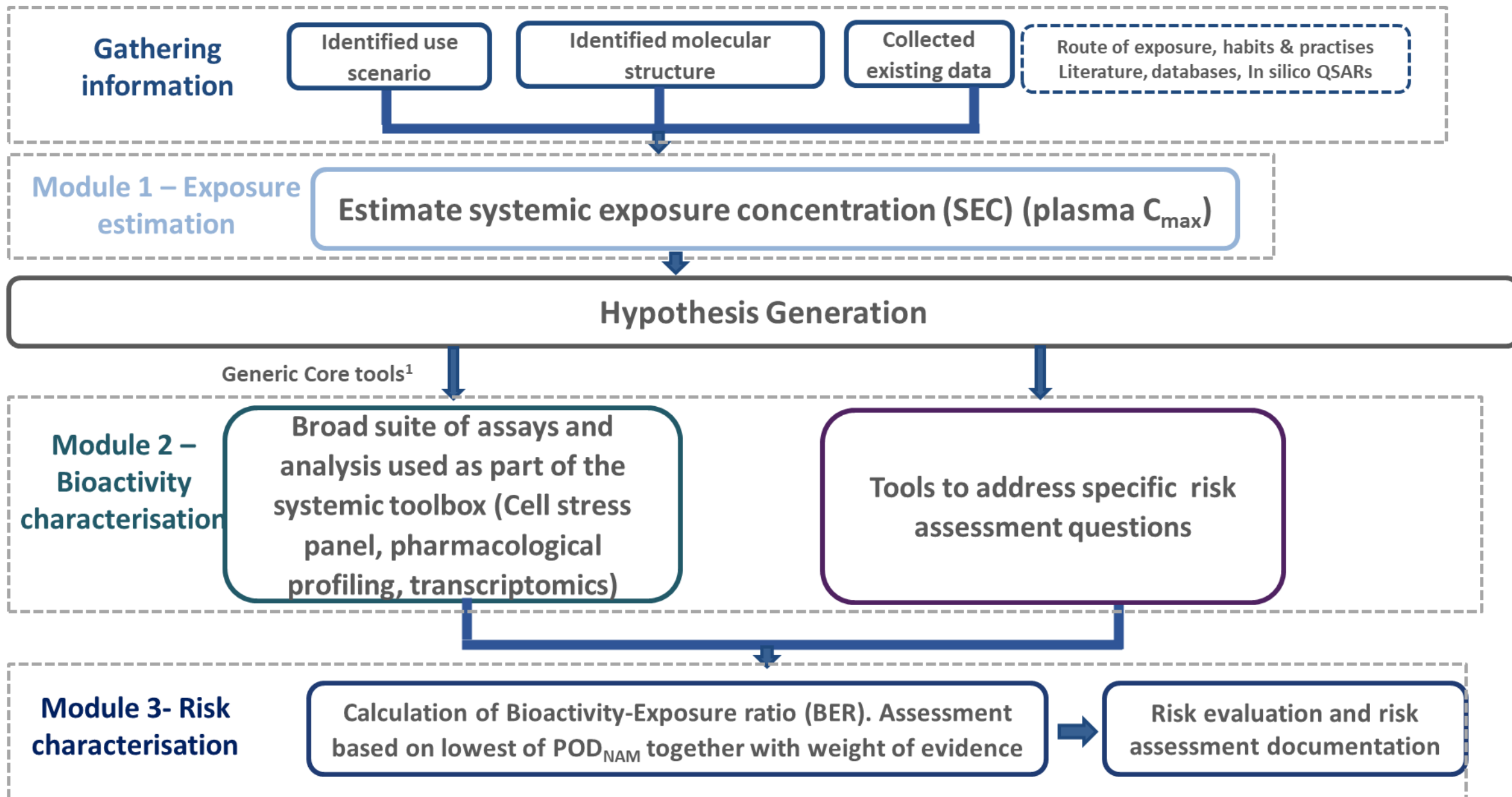
Exposure models (PBK, free/total concentration)



Exposure estimation: Plasma C_{max} , organ distribution, AUC



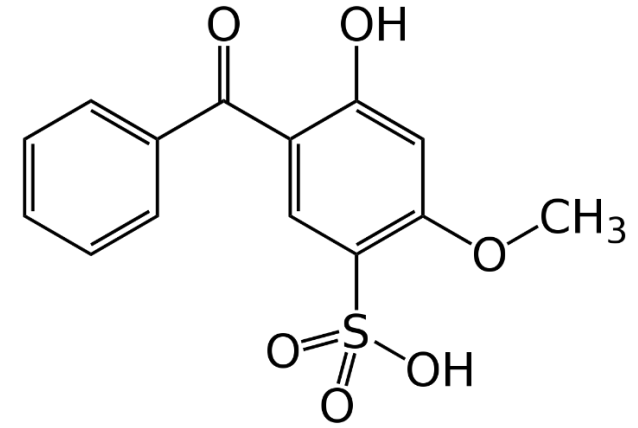
Overall approach for Benzophenone-4 (BP-4)



¹Middleton et al. (2022) *Toxicol Sci* (<https://doi.org/10.1093/toxsci/kfac068>)

Gathering information: Use scenario and molecular structure

- Benzophenone-4 (CAS No. 4065-45-6; EC No. 223-772-2) **has been used up to 5% in Europe in cosmetics for decades as an ultraviolet (UV) filter** and provides protection of the skin and hair from the harmful effects of the sun.
- Benzophenone-4 is **water soluble**, given the presence of a sulphate group in its chemical structure (see section 3.1.1), **and an anion at physiological pH**
- It is also used as a **product protectant at much lower % inclusion levels** as a UV stabiliser protecting cosmetic formulations against chemical breakdown by sunlight
- **The specific use scenario** of this case study is for **dermal application of a leave-on sunscreen body lotion product** containing **benzophenone-4 at 5% w/w**



Daily use of sunscreen lotion UV-filter*:

- Amount of sunscreen applied = 18 g/day divided into two applications of 9g (SCCS recommendation)
- External dose= 15 mg/kg bw/day

Gathering information: in silico tools

In silico tools for toxicity endpoints: OECD QSAR TOOLBOX, TOXTREE, DEREK NEXUS and DEREK METEOR (metabolism)

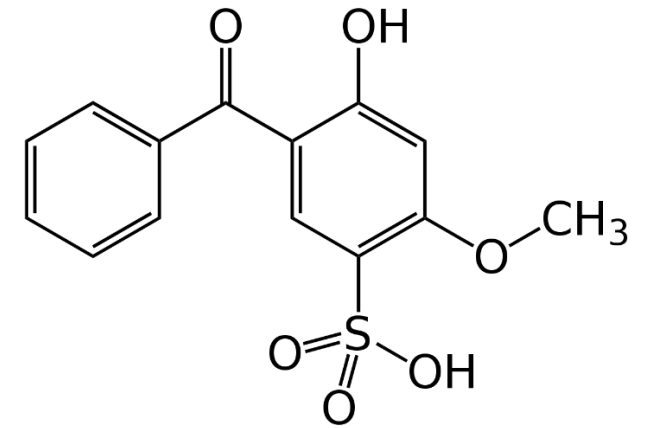
- **Benzophenone-4 did not trigger many alerts within the tools used.**

The most common alert across the tools was for skin sensitisation, or protein binding as an indication of skin sensitisation, in the DEREK, TIMES and OECD Toolbox outputs.

- **Benzophenone-4 triggered one potential alert for estrogen receptor binding in the VEGA profiler, however this was not consistent across other profilers that also assess estrogen receptor activity.**

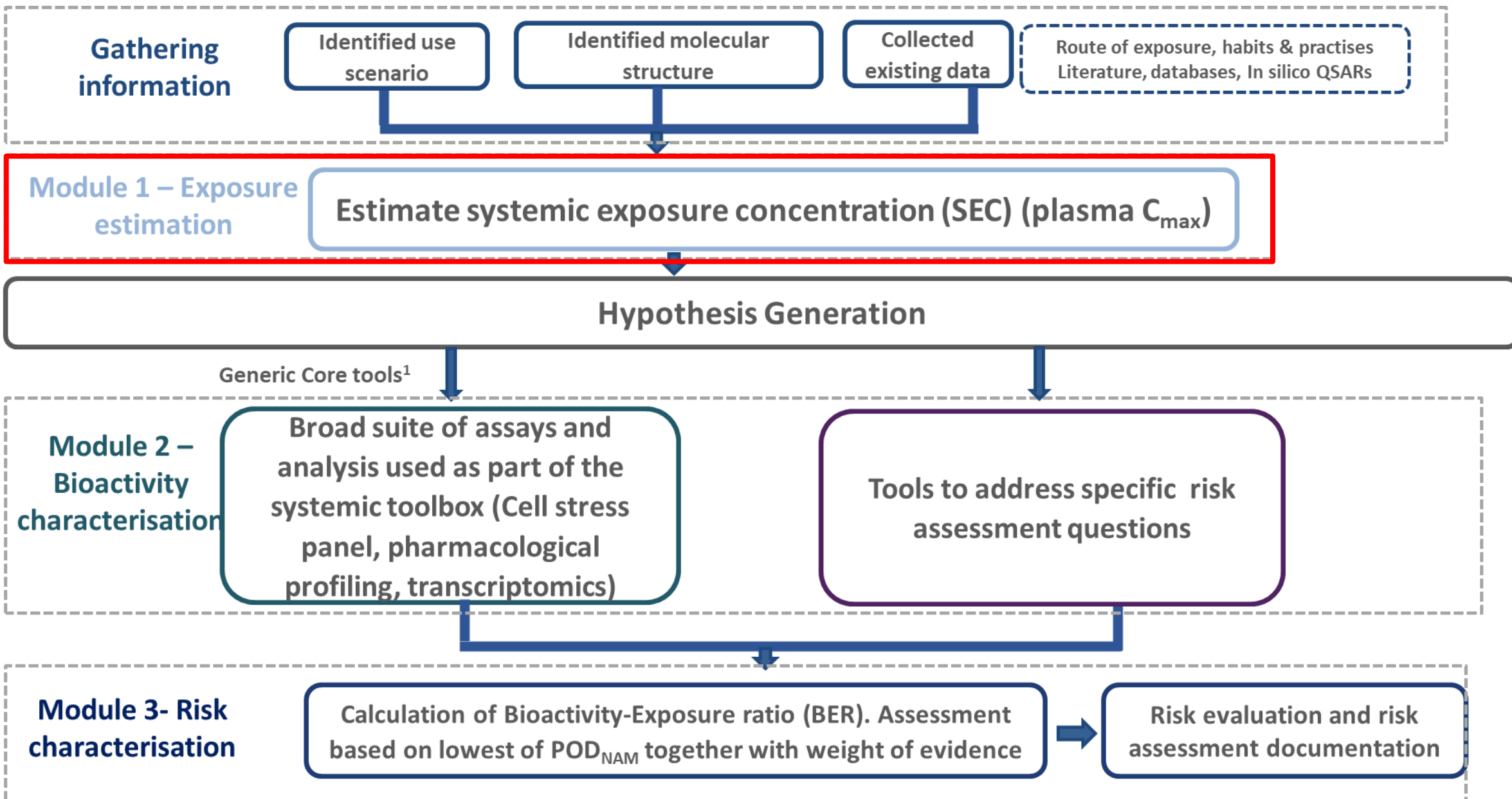


Follow up with in vitro assays to confirm whether or not BP-4 binds to estrogen receptor and other endocrine related endpoints – CALUX EATS (estrogenic, androgenic, thyroidogenic and steroidogenesis)



CAS No. 4065-45-6; EC No. 223-772-2; sulisobenzone; 2-Hydroxy-4-methoxybenzophenone-5-sulphonic acid)

Overall approach for Benzophenone-4 (BP-4)



¹Middleton et al. (2022) *Toxicol Sci* (<https://doi.org/10.1093/toxsci/kfac068>)

Module 1: Exposure assessment

From applied dose to internal concentrations

External dose

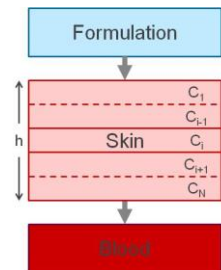
- Route of exposure
- Consumer use (Habits & Practices)
- Applied dose (external concentration)



ADME parameters

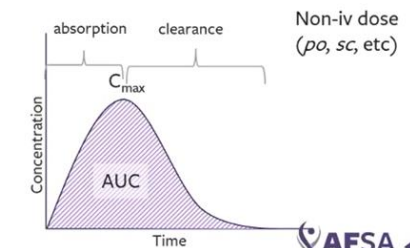
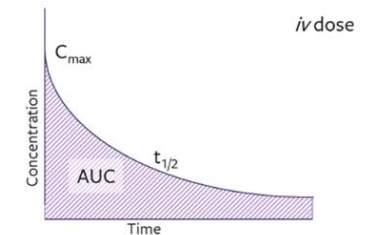
Absorption
Distribution
Metabolism
Elimination

- Skin penetration
- Phys-chem properties
- Hepatic clearance
- Fraction unbound
- Blood:plasma ratio



Kinetic profile of chemical

Physiologically-based kinetic (PBK) modelling
– Internal concentration (plasma, urine, organ-level)



Images from: AFSA training module "Dosimetry (Internal Exposure)", 2022

PBK modelling inputs– Exposure scenario, target individual/population, ADME parameters

Exposure scenario

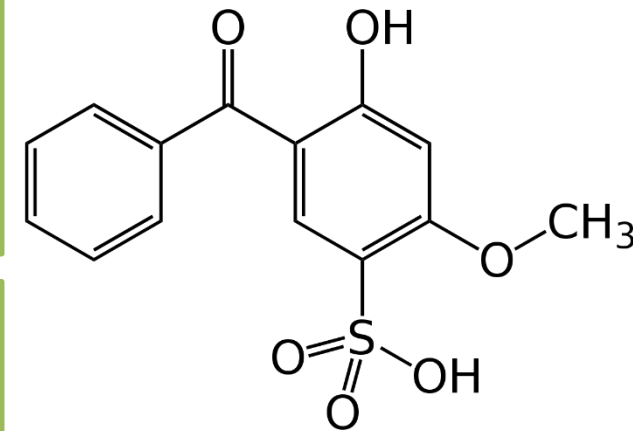
- 5% in Sunscreen product,
- 18g/day, two times, 9g/application,
- On body and face 17500cm² (total body area)

Physiological parameters

- Adult female, 30 years old, 60 kg (SCCS NoG 12th revision)
- PEAR (Population Estimates for Age-Related -Physiology™) was used to calculate organ weights, volumes, perfusions, and tissue-plasma partition coefficients for the 30 year old, 60 kg bodyweight person.

ADME data generation – in vitro

- Dermal absorption (OECD TG
- Blood to plasma ratio
- Plasma protein binding
- Metabolic stability (cryopreserved primary human hepatocytes)



PBK modelling inputs – ADME results

Main observations:

In silico

- BP-4 was predicted to be cleared via liver metabolism (ECCS classification, Varma et al 2015)
- BP-4 is predicted to be substrate of several transporters by ADMET predictor

Experimental

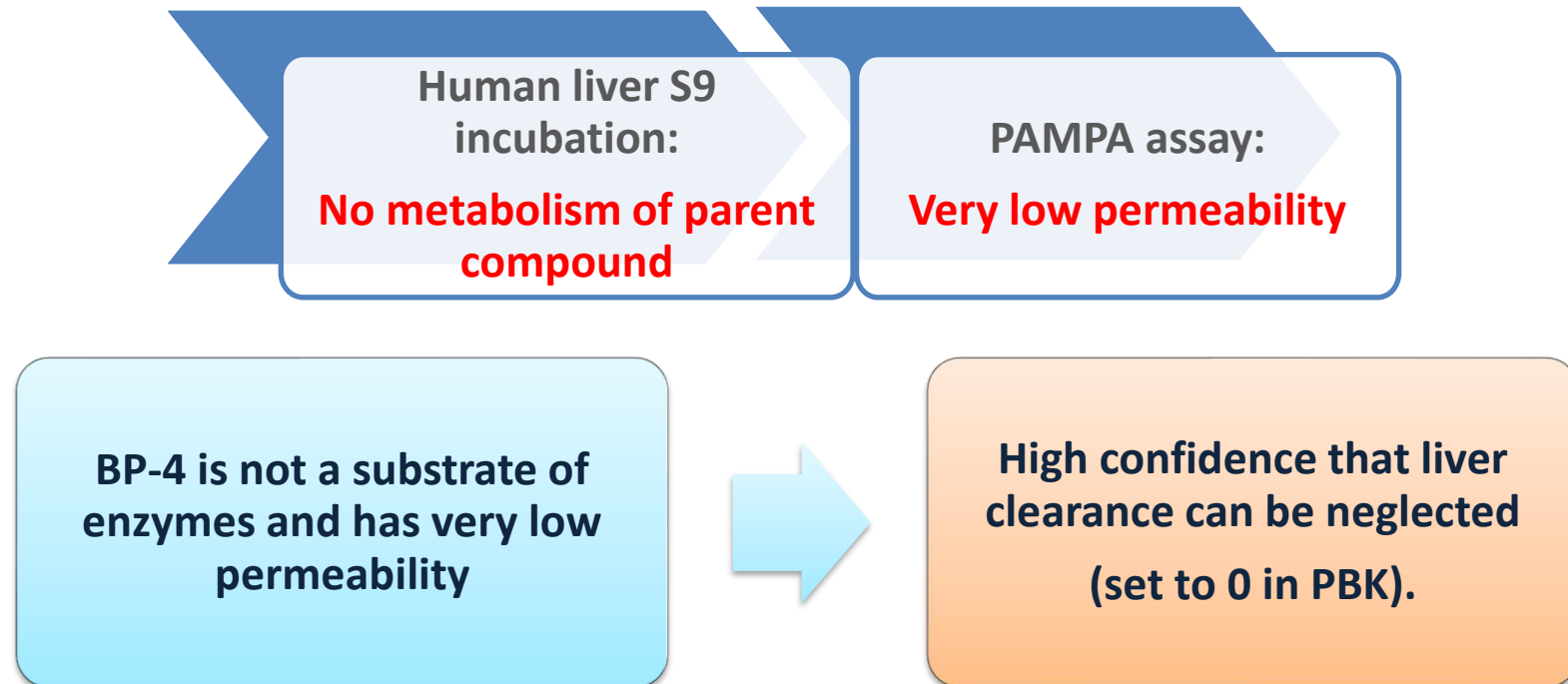
- Very low skin penetration
- BP-4 stable in human hepatocytes. Hepatic intrinsic clearance <2.5L/h (Below LOQ)

Conclusion: Conflicting data between in silico and experimental

Back to problem formulation - Two hypotheses:

- 1) Benzophenone-4 is not a substrate of CYP enzymes – need to confirm with a second assay using S9 fraction
 - **Note**, BP-4 is an hydrophilic compound already
- 2) Benzophenone-4 has low membrane permeability– Parallel artificial membrane permeability (PAMPA) assay

Follow up assays



Understanding chemical organ distribution and renal clearance: Is BP-4 actively transported by active transporters in kidney?

Two experimental approaches:

1. Transporter studies in transfected kidney cells in two different assays (uptake assay and vesicular assay)

2. Investigate the transport profile in kidney where all the active transporters are present and functional (freshly isolated kidney proximal tubule cells monolayer (aProximate™)).

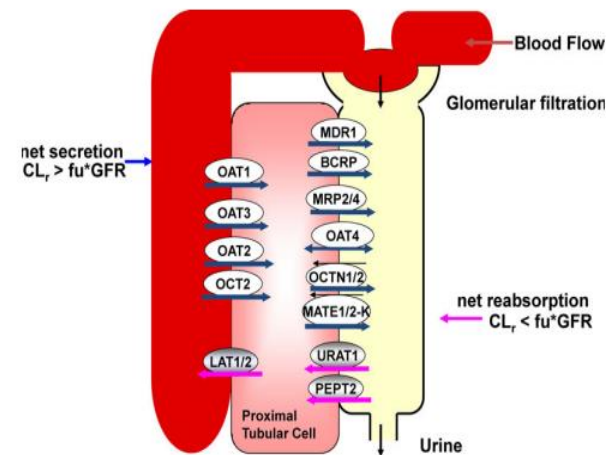
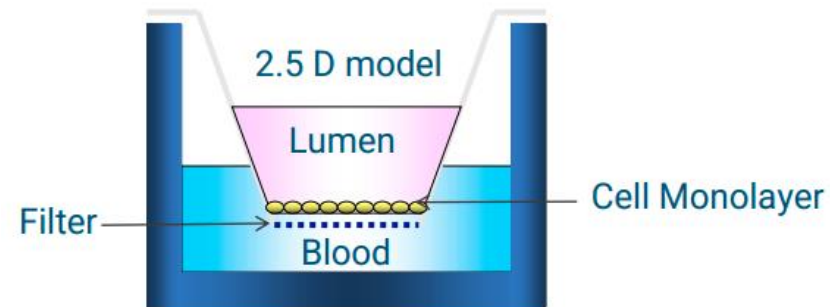


Figure 1. Mechanism of drug elimination and major transporters in the kidney. Drug elimination in the kidney is through glomerular filtration, secretion, and reabsorption process. Major transporters localized in the proximal tubule cells are depicted. The blue arrows indicate secretion, and the pink arrows indicate reabsorption.



B-A → blood to urine → active secretion
A-B → urine to blood → reabsorption

Newcells aProximate™ platform

<https://doi.org/10.1002/jcph.702>

Understanding chemical organ distribution and renal clearance: Is BP-4 actively transported by active transporters in kidney?

1. Transporter studies in transfected kidney cells in two different assays (uptake assay and vesicular assay)

Results:

- Substrate of the **influx transporters**, OAT1, OAT2, OAT3 and a substrate of the **efflux transporters**, BCRP and MRP4.
- All these transporters are **expressed in the kidney**, although OAT-2, BCRP and MRP4 are expressed both in **kidney and liver**

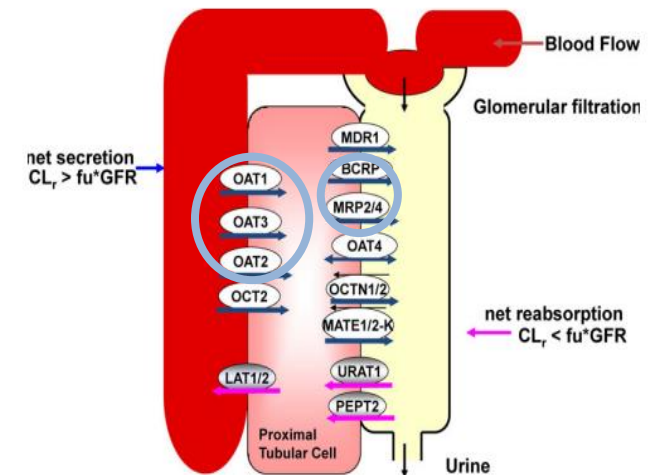
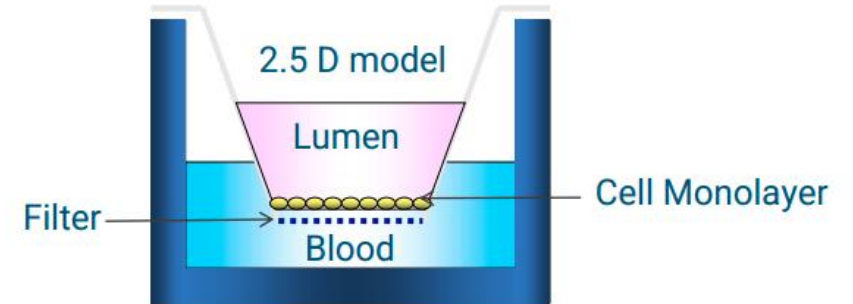


Figure 1. Mechanism of drug elimination and major transporters in the kidney. Drug elimination in the kidney is through glomerular filtration, secretion, and reabsorption process. Major transporters localized in the proximal tubule cells are depicted. The blue arrows indicate secretion, and the pink arrows indicate reabsorption.

Understanding chemical organ distribution and renal clearance: Is BP-4 actively transported by active transporters in kidney?

2. Investigate the transport profile in kidney where all the active transporters are present and functional (freshly isolated kidney proximal tubule cells monolayer (aProximate™)).



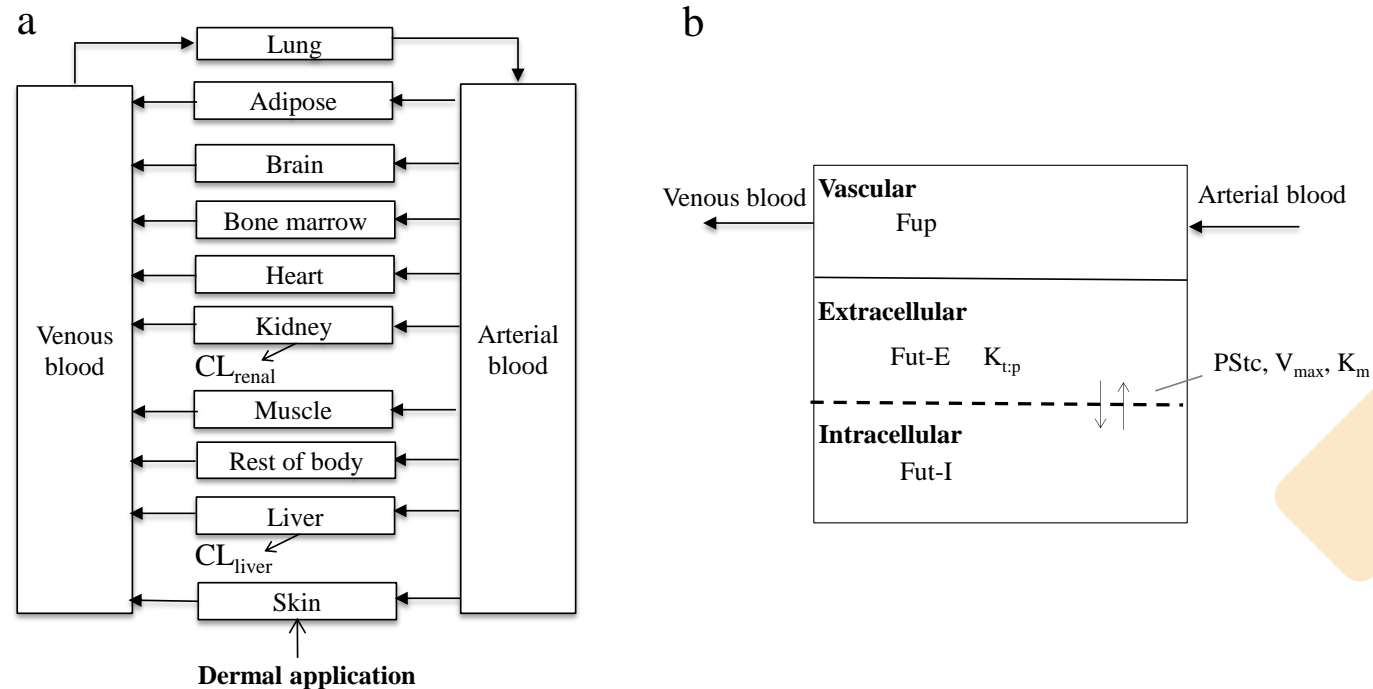
B-A → blood to urine → active secretion
A-B → urine to blood → reabsorption

Results:

- Transport in the proximal tubule cells is **equally efficient in both directions leading to no net movement**
- However, donor variability has been observed that in **1 donor, active secretion was shown to be the main excretion** route at biologically relevant concentrations

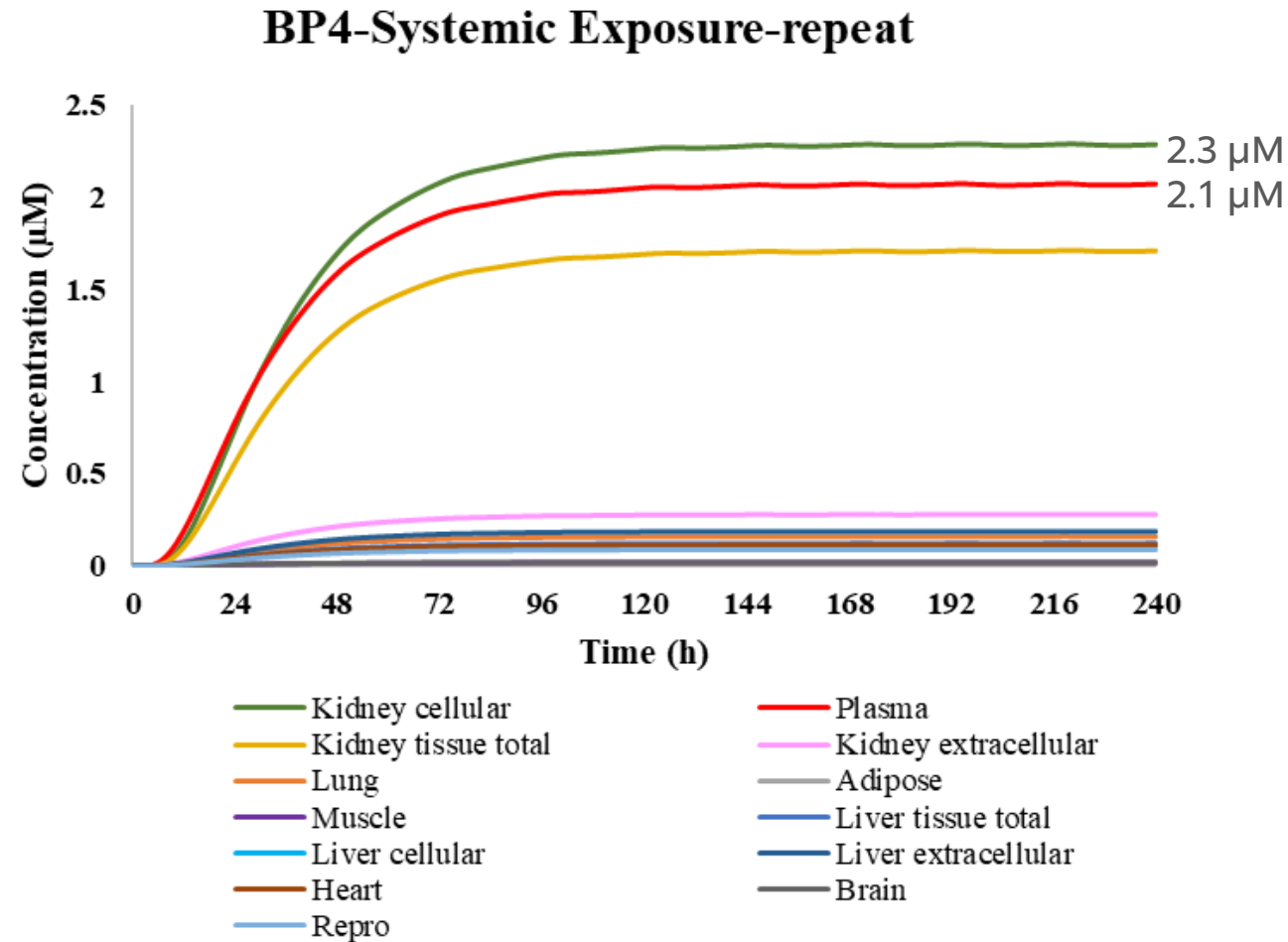
Update PBK model

- Set BP-4's distribution to each compartment to be modelled as permeability-limited
- Liver clearance set to 0
- Active transport in the liver was modelled by incorporating kinetic parameters for the transporters (OAT-2, BCRP and MRP4).
- $GFR \cdot F_{up}$ was used to calculate renal excretion of benzophenone-4, accounting for filtration only to be conservative



Human PBK model structure for BP4

Internal concentration: Deterministic PBK model simulation of C_{max} for an adult female (30 years old, 60 kg)



Benzophenone-4 concentrations in plasma and different tissues after repeated exposure of body lotion 18g/day, i.e., 9g twice per day for a period of 10 days, with 5% benzophenone-4, on the whole body.

How to address uncertainty in the PBK modelling?

We personally care

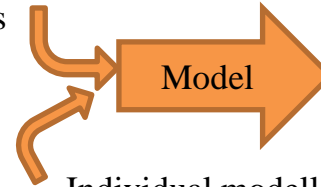


Cosmetics Europe
the personal care association

Strategies in addressing uncertainty in PBK estimation

Deterministic PBK modelling

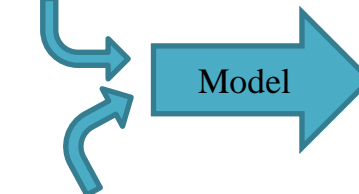
Point estimate values for input parameters



Individual modelled (30 year-old 60 kg female, European)

Probabilistic population PBK modelling

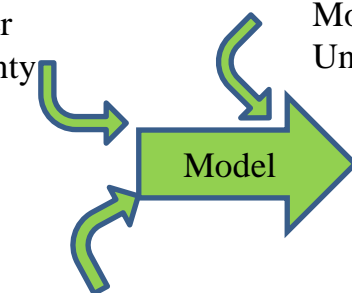
Parameter Uncertainty ('informed' distribution for the most sensitive parameters)



Population Variability

Probabilistic population PBK+CMED modelling

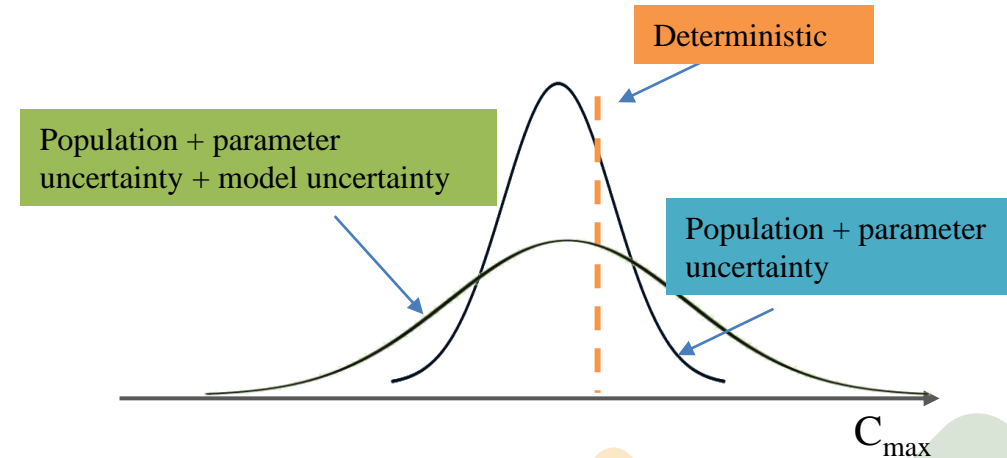
Parameter Uncertainty



Population Variability

Model Uncertainty

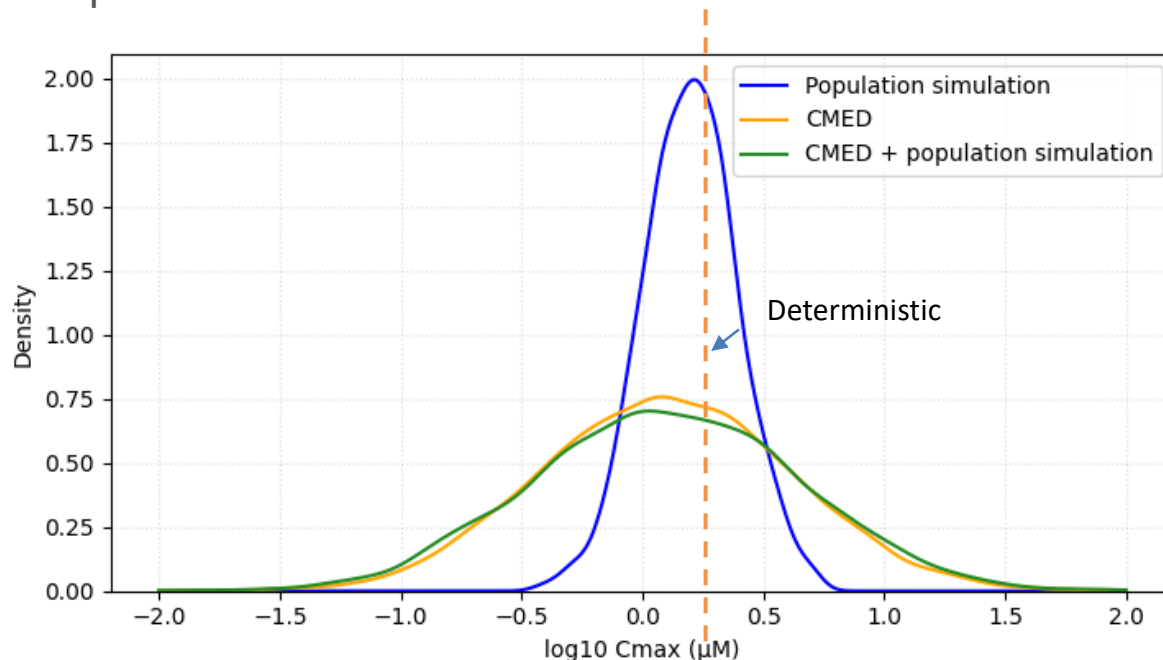
Predicted C_{max} based on different approaches characterising uncertainty



Probabilistic PBK modelling + CMED model to account for population, parameter and model uncertainty

To account unknown-unknowns e.g. model uncertainty

- C_{max} Error Distribution (CMED): A complementary approach to characterise PBK prediction uncertainty as published in Middleton *et al.* 2022.
- This model seeks to quantify the error distribution of estimates of plasma C_{max} by looking at the difference between PBK predictions of C_{max} and existing measured values in human clinicals for several exposure scenarios.
- This model can be used to estimate the distribution of the possible prediction errors for future chemical and exposure scenario.



Deterministic PBK model for adult female 60 kg	Distribution of C_{max} within the in vivo population estimated by combining CMED model and GastroPlus™ population simulation (µM) (green curve)	
Plasma C_{max} point estimate	Median (95% interval)	95 th percentile
2.1	1.3 (0.11, 15)	9.8

To summarize BP-4's kinetic behavior in the human body:

- Overall, upon dermal absorption only a small amount of BP-4 enters systemic circulation, after which BP-4 remains unchanged due to negligible liver clearance.
- It has low tissue distribution due to low partitioning and limited passive diffusion of cell membranes (charged at physiological pH).
- It can be taken up into the kidney and then excreted to urine via active transport and can be reabsorbed back to into the bloodstream, however due to no preferred direction of movement glomerular filtration determines the overall renal excretion rate.
- BP-4 can also distribute into the liver.
- Successive doses result in accumulating concentrations of BP-4 in the body until a steady state is reached at around 100h when there is an equilibrium reached between the low absorption and low excretion into the urine.

Assessing the confidence level

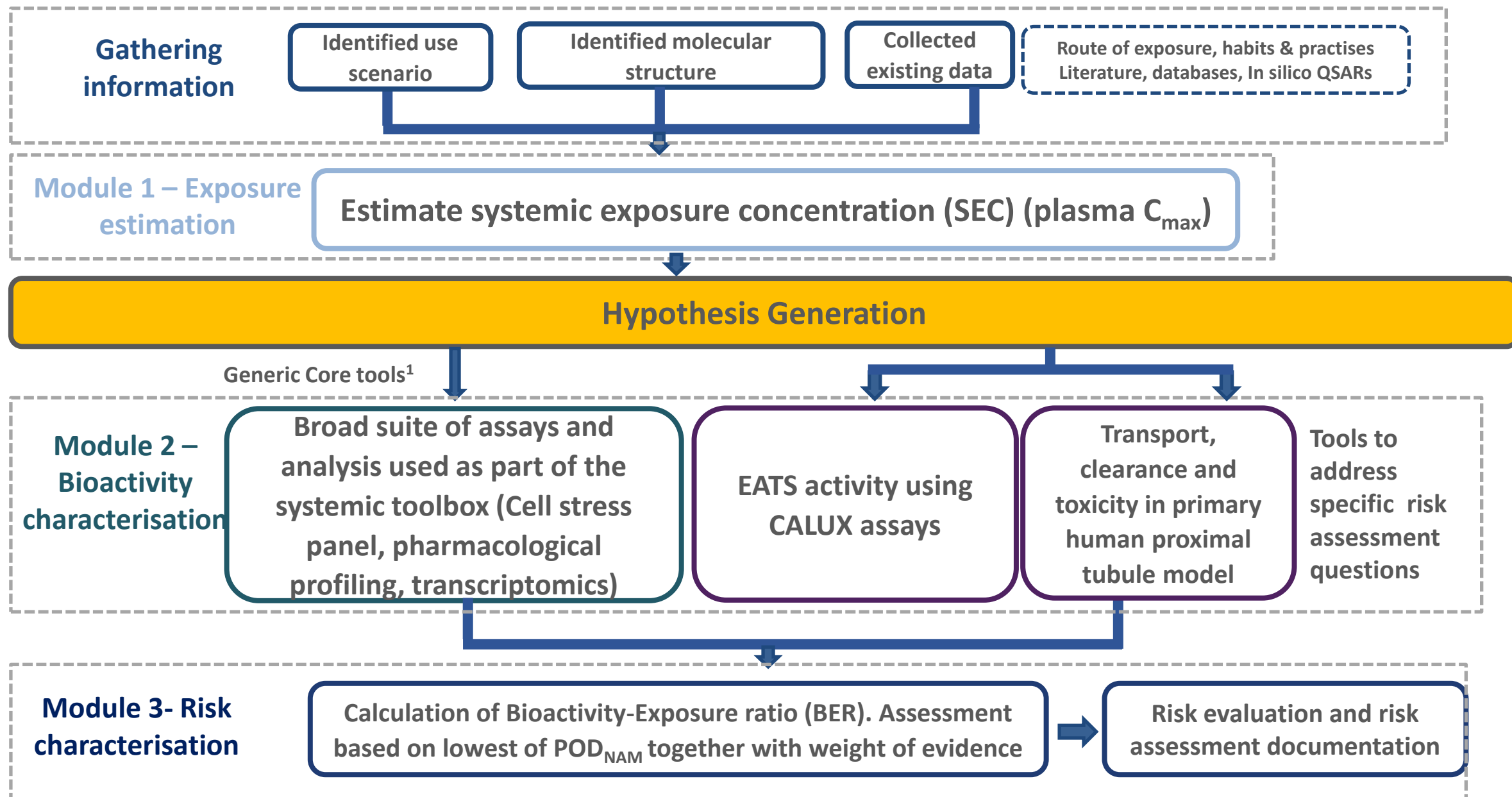
WHO questions for assessing the level of confidence in the BP-4 PBK modeling

Model evaluation aspect	level of confidence (towards the accuracy)	level of confidence (towards the conservatism)
Do the model structure and parameters have a reasonable biological basis?	High	High
How well does the PBK model reproduce the chemical-specific PK data under various experimental or exposure conditions?	Low	High
How reliable is the PBK model with regard to its predictions of dose metrics relevant to risk assessment?	High	High

Conclusions

- ✓ The stepwise way of data generation and refinement, using relevant and robust approaches for parameter determination, support the reliability of input parameters and provide a sound biological basis for the model structure.
- ✓ Although human clinical data are not available for validation, the sensitivity and uncertainty analyses and the probabilistic modelling performed provided assurance that the predictions are fit for purpose and provides conservative estimates of human systemic exposure.

Overall approach for Benzophenone-4 (BP-4)



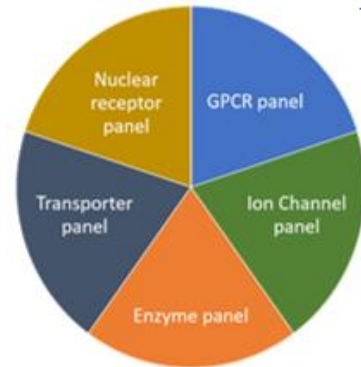
¹Middleton et al. (2022) *Toxicol Sci* (<https://doi.org/10.1093/toxsci/kfac068>)

Hypothesis Generation

- 1) **Biological activity measured using a broad suite of human-relevant test systems is sufficiently protective. If bioactivity is not observed at concentrations experienced systemically in consumers then there are no adverse effects.**
- 2) In silico tools predicted binding to estrogen receptor.
- 3) PBK model indicated that the concentration of BP-4 is higher in the kidney than in any other organ, therefore a relevant kidney cell model was included in the testing strategy.

Module 2: Broad suite of assays and analysis used as part of the systemic toolbox

In vitro pharmacological profiling



~79 targets

eurolins | Cerep



Bowes et al. 2012. Nat Rev Drug Discov 11(12): 909-22

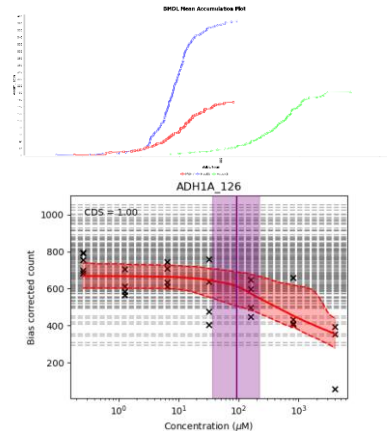
Transcriptomics was applied as a broad nontargeted biological screen

To investigate possible interactions between coumarin and the 44 key targets involved in drug attrition – now expanded to 79 targets

To characterize non-specific biological activity which is not mediated via a specific protein/receptor interaction

High-Throughput transcriptomics (HTTr)

- TempO-seek 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

Cell stress panel (CSP)

- 36 biomarkers covering 10 cell stress pathways
- HepG2
- 24hr exposure
- 8 concentrations
- Dose-response analysis using BIFROST model

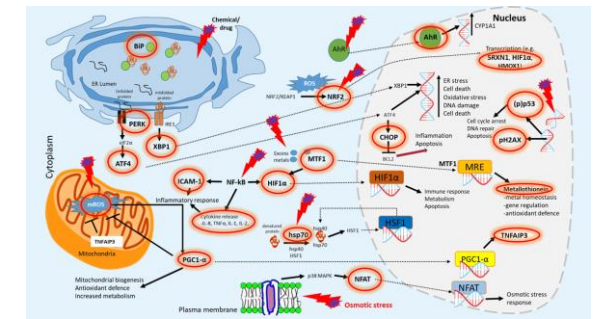


Image kindly provided by Paul Walker (Cyprotex)

Hatherell et al. 2020. Toxicol Sci 176(1): 11-33

Hypothesis Generation

- 1) Biological activity measured using a broad suite of human-relevant test systems is sufficiently protective. If bioactivity is not observed at concentrations experienced systemically in consumers then there are no adverse effects.
- 2) **In silico tools predicted binding to estrogen receptor.**
- 3) **PBK model indicated that the concentration of BP-4 is higher in the kidney than in any other organ, therefore a relevant kidney cell model was included in the testing strategy.**

Module 2: Tools to address specific risk assessment questions

2. In silico prediction for estrogen binding



EATS activity: estrogenic, androgenic, thyroidogenic and steroidogenesis

- **CALUX bioassays and binding assays: TTR-TR β - and hTPO**
- **U2-OS** incorporating the firefly **luciferase reporter gene** coupled to Responsive Elements (REs)
- **12 concentrations.** Calculation of AC50, LOEC and NOEC

3. Benzophenone-4 concentration was predicted to be higher in the kidney than any other organ

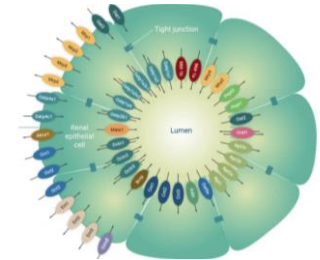
4. Cell models in the toolbox have limited expression of the relevant transporters



Renal Toxicity

Renal biomarkers (3 donors, duplicate per donor), 8 concentrations, 24h and 72h timepoints:

- KIM-1
- NGAL
- Clusterin
- TEER (Day 0 and Day 3)
- ATP
- LDH
- Toxicogenomics (3 donors, 2 duplicates per donor), 8 concentrations, 24h and 72h timepoints
- Omeprazole and cisplatin added as benchmarks/positive controls



Newcells aProximate™ platform

Piyush Bajaj et al. 2020. Toxicology. 442, 152535

Results from the key NAMs- Deriving Points of Departure (PoDs)

In vitro Pharmacological profiling

- Tested up to 10 μM
- ~83 targets compiled by Cosmetics Europe Safety pharmacology WG
- **No hits**

Calux assays

- No agonism or antagonism of ER, AR or TR and no effect on production of oestrogens or androgens $\pm\text{S9}$
- Activity towards hTPO and TTR was found at high concentrations (LOEC= 300-600 μM).

Cell Stress Panel

- Global $\text{POD}_{\text{NAM}} = 140 \mu\text{M}$

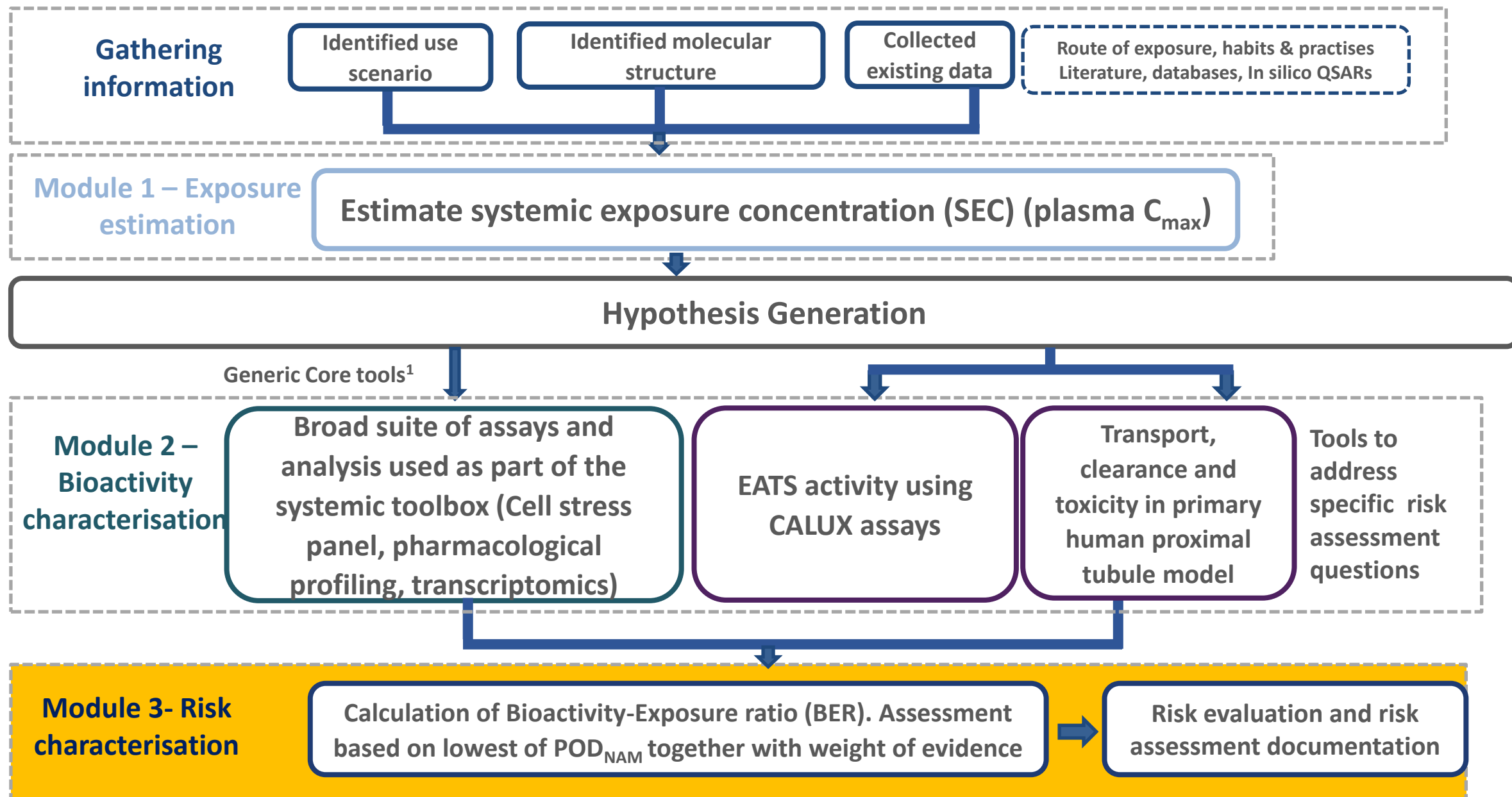
HTTr (HepG2, HepaRG, MCF7, PTC)

- Two approaches to calculating POD – BIFROST (gene level) and BMDL (pathway level); PODs varied from 4.2 – 530 μM

Renal biomarkers (PTC)

- No significant response for BP-4 (Cisplatin and Omeprazole gave expected dose-response at 72-h)

Overall approach for Benzophenone-4 (BP-4)



Bioactivity: exposure ratio calculation 1/2

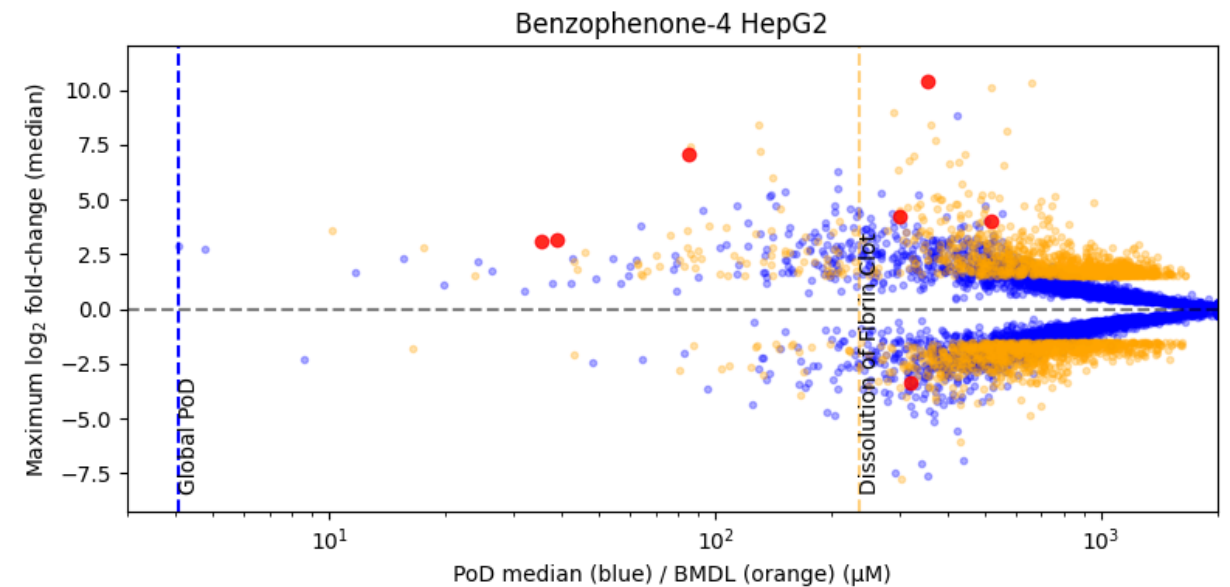
NAM	Cell type	POD _{NAM} Type	POD _{NAM} Value (µM)	BER (using C _{max} of 2.1 µM)	BER from individual C _{max} (µM) (CMED + PBK population simulation)	
					Median (95% interval)	Prob. BER>1
Cell stress panel	HepG2	Global PoD	140	67	110 (11, 1200)	1.0
HTTr	HepG2	Global PoD	4.2	2	3.4 (0.32, 35)	0.85
HTTr	HepaRG	Global PoD	52	25	42 (4, 430)	1.0
HTTr	MCF7	Global PoD	5.5	2.6	4.4 (0.42, 45)	0.90
HTTr	HepaRG	Lowest pathway BMDL	530	252	430 (41, 4400)	1.0
HTTr	HepG2	Lowest pathway BMDL	240	114	190 (18, 2000)	1.0
HTTr	MCF7	Lowest pathway BMDL	330	157	260 (25, 2700)	1.0

Bioactivity: exposure ratio calculation 2/2

NAM	Cell type	POD _{NAM} Type	POD _{NAM} Value (µM)	BER (using C _{max} of 2.1 µM)	BER from individual C _{max} (µM) (CMED + PBK population simulation)	
					Median (95% interval)	Prob. BER>1
Calux (hTPO-inhibition)	-	LOEC	300	143	240 (23, 2500)	1.0
Calux (T4 binding to TTR)	-	LOEC	630	300	510 (48, 5200)	1.0
Renal biomarkers (24 hr exposure)	PTC	Global PoD	>1000	NA	NA	NA
Renal biomarkers (72 hr exposure)	PTC	Global PoD	>1000	NA	NA	NA
HTTr (renal cells) (24 hr exposure)	PTC	Global PoD	320	152	260 (25, 2600)	1.0
HTTr (renal cells) (72 hr exposure)	PTC	Global PoD	320	152	260 (25, 2600)	1.0

Safety assessment discussion

- Lowest BER across all PODs was obtained from HTTr in HepG2 cells when the BIFROST method was used (POD of 4.2 μM ; deterministic BER of 2)
 - *Single gene change of CYP 1A1 – vertical dashed blue line*
 - *Lowest BMDL in the same cell line is 240 μM (average of red circles)*
- This provides some assurance that the gene changes seen at 4.1 μM may be of limited toxicological significance.



Maximum fold-change in expression against BIFROST probe-level median POD (blue), and BMDEExpress2 probe-level BMDLs (orange). Global POD calculated by BIFROST model (blue dotted line) and minimum pathway BMDL obtained from BMDEExpress2 (orange dotted line). Red circles are the BMDExpress2 probe-level BMDLs contributing to the lowest pathway average. Global POD = CYP1A1 probe

A qualitative consideration of uncertainty for the safety assessment of benzophenone-4 at 5% in sunscreens (Table 15 in the Dossier)

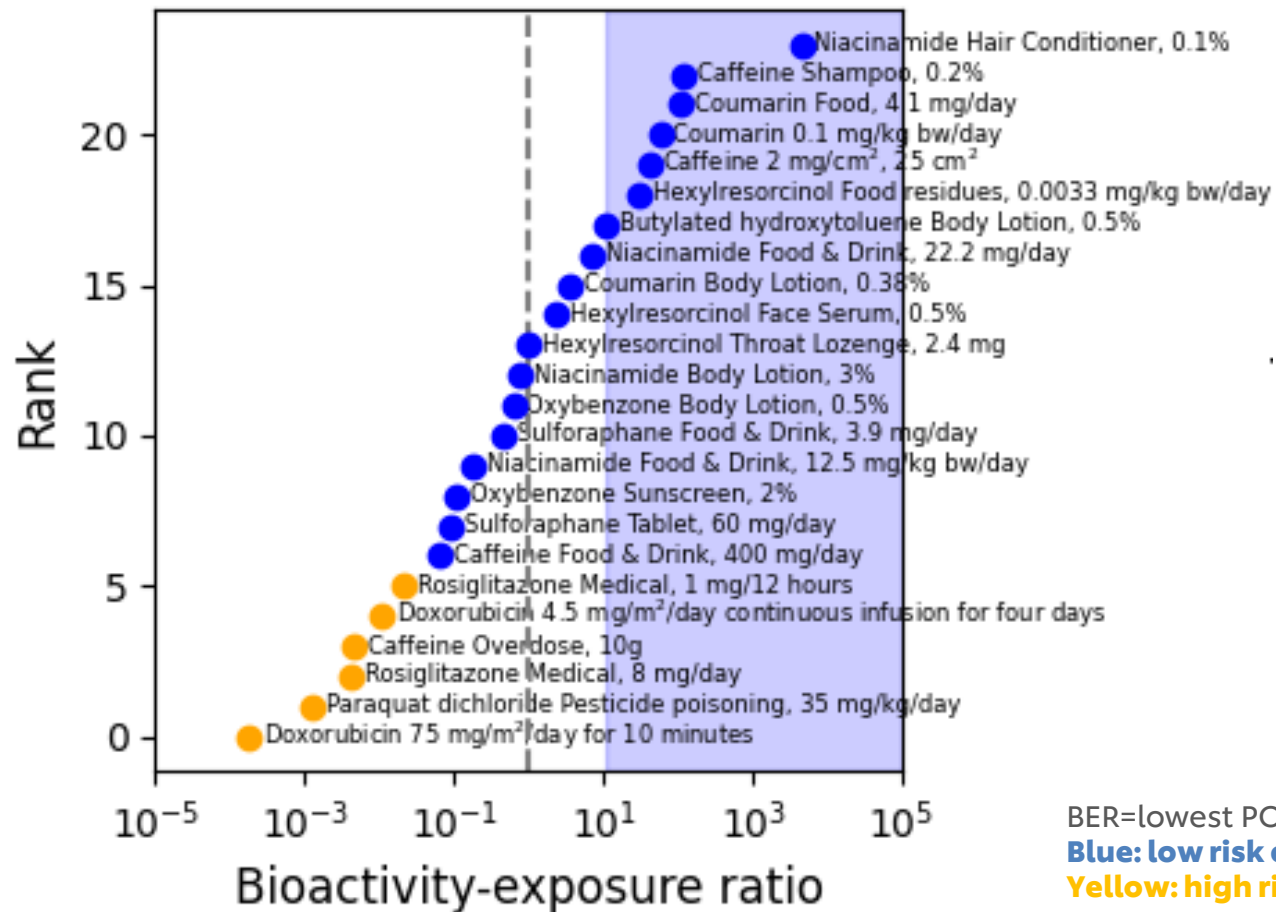
Area	Level of certainty (rationale)	Is value likely to be an over- or under-estimate (rationale)	Impact on risk assessment decision
Range of biomarkers assessed	Moderate (There is increasing evidence that POD_{NAM} obtained from the core NAMs, IPP, CSP and HTTr are protective for a range of chemicals (Middleton <i>et al.</i> , 2022) and previous case studies (Baltazar <i>et al.</i> , 2020, OECD phenoxyethanol). The hypothesis and exposure driven approach led to the inclusion of additional NAMs to investigate potential endocrine activity and kidney toxicity)	Given the low activity of benzophenone-4 across all available assays together with its kinetic profile (low passive permeability and low organ distribution) it is considered unlikely a specific MoA exists that would affect the safety assessment	There are remaining uncertainties regarding the protectiveness of the tools utilised for a broader range of chemistries. Confidence could be increased by assessing how protective the range of biomarkers are for many more compounds and whether different biomarkers are needed to ensure the <i>in vitro</i> PoD is protective compared with the <i>in vivo</i> PoD
Use of short-term tests <i>in vitro</i> to inform about risks of long-term human exposure	Moderate (There is increasing evidence showing that for many chemicals use of a short-term <i>in vitro</i> PoD, and short-term transcriptomics data in particular are protective for long-term target organ effects (Thomas <i>et al.</i> , 2013; Paul Friedman <i>et al.</i> , 2020; Middleton <i>et al.</i> , 2022) and previous case studies (Baltazar <i>et al.</i> , 2020, OECD phenoxyethanol))	Previous analyses show that short term <i>in vitro</i> PoDs are protective in most cases	No correction made for duration of exposure from <i>in vitro</i> to human exposure.

But...from a quantitative perspective... How do we define an acceptable BER to conclude an exposure to a give chemical is low risk?

Conceptually, with the following assumptions a $BER > 1$ indicates a low risk of adverse effects in consumers following use of the product:

1. The in vitro measures of bioactivity provide appropriate biological coverage
2. There is confidence that the test systems are at least as sensitive to perturbation as human cells in vivo
3. The exposure estimate is conservative for the exposed population

Building confidence in a NAM toolbox – same core toolbox (PBK, HTTR, Cell stress panel, IPP) was applied to 10 chemicals and 24 exposure scenarios



- Based on our first evaluation -> Conclude low risk at PBK L2 if the BER point estimate >11 (blue shaded region)
- NAM Systemic toolbox 100% protective for high-risk chemical exposure scenarios
- Very conservative safety decisions using Tier 1 toolbox alone

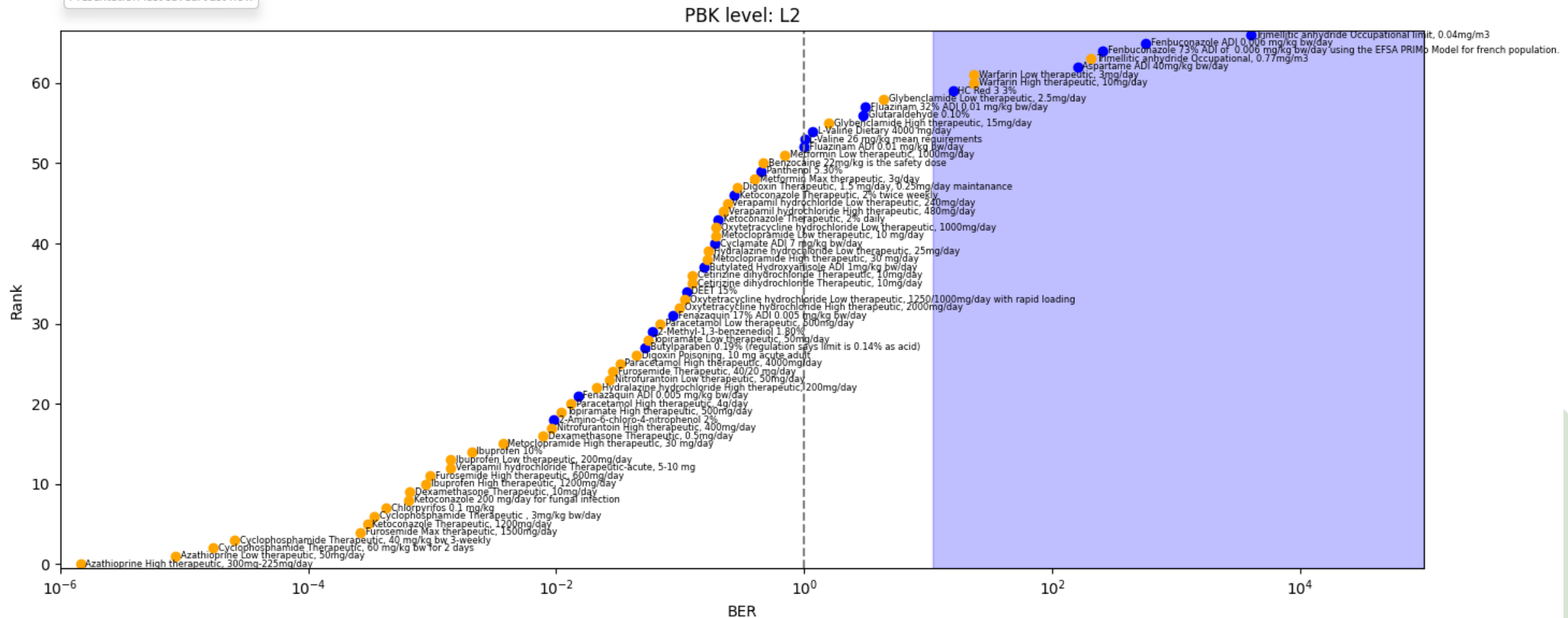
BER=lowest POD/Plasma Cmax

Blue: low risk chemical-exposure scenario

Yellow: high risk chemical-exposure scenario

Blue shaded region BER > 11

NAM Systemic toolbox remains protective (93%) when 38 additional chemicals and 70 exposure scenarios were tested (manuscript in preparation) using the previous BER thresholds



Risk assessment conclusion for BP-4

- The lowest BMDL PODNAM was calculated to be 240 μM (HepG2 cells), that results in a BER of 114, significantly greater than the BER of 2 derived from the same cell type using the gene-level global PODNAM from the BIFROST method.
- This provides some assurance that the gene changes seen at 4.1 μM may be of limited toxicological significance.
- The BER calculated from the deterministic Cmax and cell stress panel global POD (the next lowest POD) was 67.
- For all other NAMs, it is very likely that the BER is above 1 for all the individuals in the population.
- Based on the tools and test systems used in this assessment and the assumptions described in the uncertainties table, **consumer internal exposures would need to be greater than those predicted to lead to toxicologically significant systemic biological activity in consumers.**

Conclusions & reflections

- Case studies have demonstrated it is possible to integrate exposure estimates and bioactivity points of departure to make a safety decision.
- These case studies showed that the approach is exposure-led and follows a tiered approach for both exposure and bioactivity
 - Bespoke NAMs can be added to the NGRA to fill gaps identified along the process
- ‘Early tier’ in vitro screening tools show promise for use in a protective rather than predictive capacity.
- NGRA requires a mindset shift and a multidisciplinary team

Acknowledgments

Matt Dent

Hequn Li

Sophie Cable

Nicky Hewitt

Beate Nicol

Joe Reynolds

Sophie Malcomber

Sharon Scott

Jade Houghton

Predrag Kukic

Andrew White

Richard Cubberley

Sandrine Spriggs

Ruth Pendlington

Katie Przybylak

Alistair Middleton

BP4 Consortium

Cosmetics Europe/LRSS Case study Leaders Team

Pharmacelsus

Eurofins

BioClavis

Cyprotex

SOLVO

BioDetection Systems

NewCells