



Full length article

Potential systemic availability classification of chemicals for safety assessment

René Geci^{a,b,*}, Alicia Paini^{a,2}, Andrew P Worth^c, Lars Kuepfer^{b,1}, Stephan Schaller^{a,1}

^a ESQlabs GmbH, Saterland, Germany

^b Institute for Systems Medicine with Focus on Organ Interaction, University Hospital RWTH Aachen, Aachen, Germany

^c European Commission, Joint Research Centre (JRC), Ispra, Italy

ARTICLE INFO

Handling Editor: Adrian Covaci

Keywords:

Physiologically based kinetic (PBK) modelling
New approach methodologies (NAMs)
Next-generation risk assessment (NGRA)
Chemical safety assessment
High-throughput PBK modelling
Toxicokinetics

ABSTRACT

The assessment of chemical safety is essential for protecting human health, yet current approaches have severe limitations. They are unable to address the rapidly growing number of substances requiring evaluation and further rely on animal testing, which the public expects to be phased out for ethical reasons. To address this, the European Partnership for Alternative Approaches to Animal Testing (EPAA) proposed the use of a novel framework for the classification of systemic toxicity based on new approach methodologies. One dimension of this framework is the grouping of substances into different classes of Potential Systemic Availability (PSA) concern (low, medium, high). But so far it remained conceptually and practically unclear how this classification can be achieved. Here, we outline the theoretical considerations for a health-protective definition of PSA concern classes and present a method for the quantitative evaluation of this property. Using high-throughput physiologically based kinetic modelling, we are able to classify the PSA concerns of 139 out of 150 EPAA NAM Designation compounds. Further, we manually annotate these compounds to evaluate the plausibility of predicted classifications against expert judgement. Our results outline under which circumstances it is appropriate to prioritise or deprioritise chemicals due to their toxicokinetic properties. However, we find that most compounds cannot be assessed on their PSA alone and need to be considered medium PSA concern so that classification into any overall systemic toxicity concern class remains possible. Future integration of bioactivity data will be necessary to fully judge the utility of our method and of the entire framework.

1. Introduction

The assessment and management of chemical safety is critical for protecting human health. But despite an extensive regulatory framework within the European Union (EU) that has conducted evaluations for decades, the European Environment Agency (EEA) estimates that only about 0.5 % of chemicals on the market have been robustly characterised, while approximately 70 % lack adequate hazard and exposure data necessary for ensuring their safe use (EEA, 2019). The primary

legislation governing industrial chemical risks in the EU is the Registration, Evaluation, Authorisation, and Restriction of Chemicals (REACH) regulation (EU, 2006). But information requirements under REACH are tied to the marketed tonnage of chemicals. For chemicals produced in quantities below 1 tonne per year, no safety information is required. Similarly, for those in the 1–10 tonnes per year category, a basic set of information is required but this is generally insufficient for conducting a comprehensive chemical safety assessment (Botham et al., 2023). Hence, many manufactured chemicals are effectively not subject

Abbreviations: ADME, Absorption, Distribution, Metabolism, and Excretion; AUC, Area under the curve; BER, Bioactivity exposure ratio; CLP, Classification, Labelling and Packaging; Cmax, Maximum concentration; EEA, European Environment Agency; EPAA, European Partnership for Alternative Approaches to Animal Testing; EU, European Union; HT-PBK, High-throughput PBK; MSAF, Mean Systemic Availability Factor; NAM, New approach methodology; PBK, Physiologically based kinetic; PBT, Persistent, bioaccumulative, toxic; PFAS, Per- and polyfluoroalkyl substances; POP, Persistent organic pollutant; PSA, Potential Systemic Availability; PSAF, Peak Systemic Availability Factor; REACH, Registration, Evaluation, Authorisation and Restriction of Chemicals; SVHC, Substance of very high concern; TD, Toxicodynamic; TK, Toxicokinetic; vPvB, very persistent, very bioaccumulative.

* Corresponding author at: Institute for Systems Medicine with Focus on Organ Interaction, University Hospital RWTH Aachen, Aachen, Germany.

E-mail address: rene.geci@esqlabs.com (R. Geci).

¹ These authors share last authorship.

² Current Address: European Food Safety Authority, Methodological and Scientific Support Unit, Via Carlo Magno 1A, 43121 Parma, Italy.

<https://doi.org/10.1016/j.envint.2025.109636>

Received 12 February 2025; Received in revised form 28 May 2025; Accepted 23 June 2025

Available online 24 June 2025

0160-4120/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

to any legal act requiring information to ensure their safety. As traditional methods have proven inadequate for addressing the vast number of unassessed chemicals, there is a need for novel approaches to chemical safety assessment.

A major limitation of current safety assessment procedures is their reliance on animal testing. Traditional approaches using animals for chemical evaluation are resource-intensive, making them impractical for assessing the increasing number of substances requiring review. Though historically a standard for toxicological evaluation, animal testing is associated with high costs and lengthy testing timelines. Additionally, there are ethical concerns and societal pressure on policymakers to phase out animal testing. For instance, six out of the ten successful European Union Citizens' Initiatives have expressed public desire for stronger protection of animals. This further underscores the need for new chemical safety assessment methods that are scientifically robust and safe but also ethically sound.

An essential part of advancing alternative safety assessment procedures is the development of methods that not only predict local but also systemic toxicity, referring to adverse effects that substances have on organs distant from the local site of contact. And while the identification and characterisation of potential hazards are central to toxicity assessment, they need to be complemented by an understanding of a substance's systemic availability to fully evaluate its potential impact on human health. This requires understanding the toxicokinetics (TK) of substances and how they are absorbed, distributed, metabolised, and excreted (ADME) within the body, as that is an essential component for determining the extent to which substances reach various body parts and whether they pose risks of adverse effects.

The NAM Designathon is an initiative by the European Partnership for Alternative Approaches to Animal Testing (EPAA) that is collaboratively exploring how new approach methodologies (NAMs) may

support the assessment of systemic toxicity within a novel regulatory framework not relying on animal testing. At the core of this initiative, a two-dimensional matrix was proposed that integrates toxicodynamic (TD) information about a compound's bioactivity with its potential TK into an overarching chemical concern classification (Fig. 1A) (Berggren and Worth, 2023). TD and TK properties are considered independently of one another and are also detached from external exposure levels to avoid reliance on exposure data, which is often unavailable or unreliable. It was suggested that under the proposed framework high-concern chemicals could be banned or heavily restricted, medium-concern chemicals could be conditionally used after application-specific risk assessments, and low-concern chemicals could be approved for unrestricted use. The aim of the initiative is to provide levels of protection equivalent to those of traditional safety assessment methods, without attempting to exactly replicate adverse effects observed in animal studies.

To challenge and explore the practical use of this framework, a list of 150 diverse chemicals was proposed by the EPAA, equally representing the three overall systemic toxicity concern classes, and participants were tasked with classifying the substances using NAMs (Supplementary File 1). The benchmark compounds further represent different use classes, such as industrial chemicals, pesticides, drugs, or substances in household products or in food (Figs. 1C–E). Except for two, all of them are small molecules.

Although the rationale behind the consideration of Potential Systemic Availability (PSA) within the NAM Designathon's assessment framework was clear, a concrete method for its evaluation had yet to be defined. Toxicity is not only influenced by a chemical's toxicokinetics but also by external exposure levels and its bioactivity (Fig. 1B). Hence, it is challenging to define a health-protective PSA classification approach that can be used independently, in the absence of information

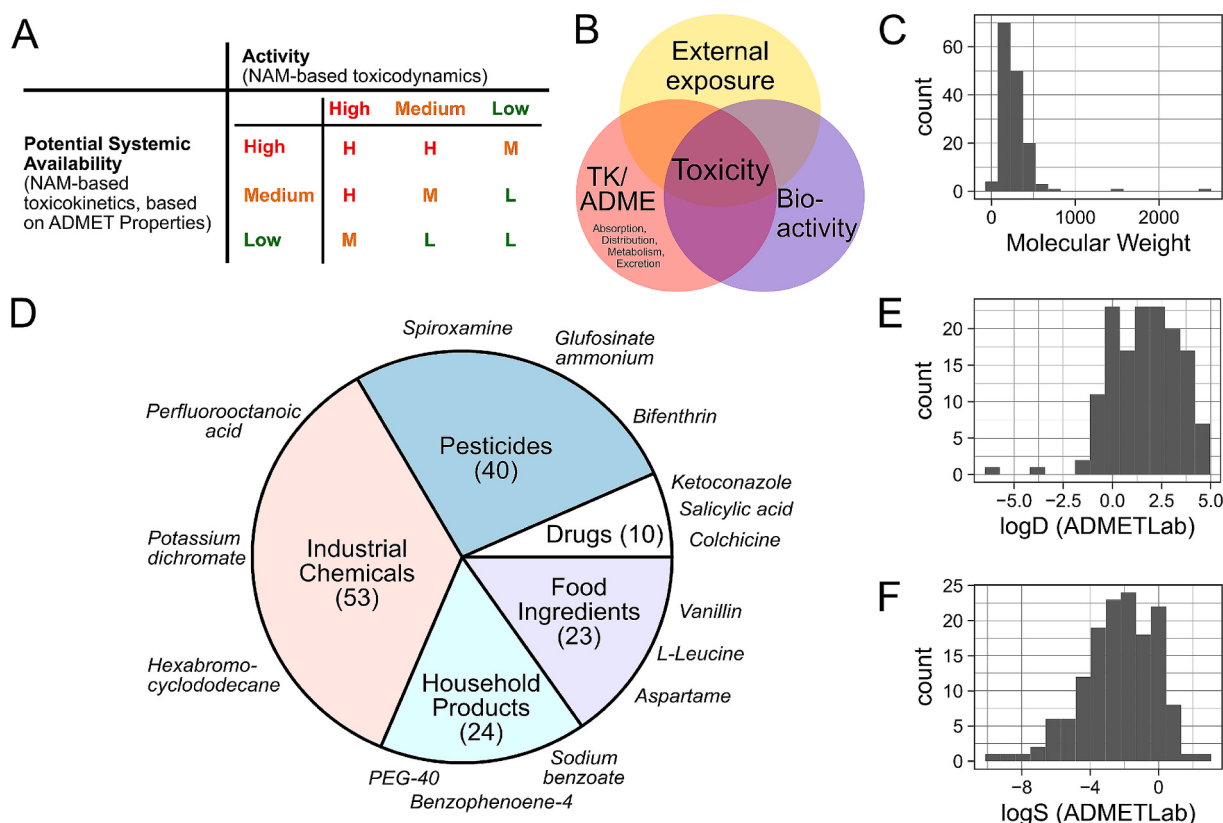


Fig. 1. (A) The EPAA NAM Designathon classification matrix integrates Activity and Potential Systemic Availability into an overall systemic toxicity concern class (adapted from Berggren and Worth, 2023). (B) Scheme illustrating that toxicity is determined by external exposure, bioactivity, and toxicokinetics. (C, E, F) Histograms showing the molecular weight, and predicted logD and logS values of NAM Designathon compounds. (D) Distribution of NAM Designathon compounds among different use classes, with three exemplary compounds of each class.

on the other toxicity determining factors. The aim of this work is to explore the scientific rationale for the classification of the PSA of chemicals into three concern classes, as it was suggested for the NAM Designation. We propose a definition of the different PSA concern classes and a practical method for efficient but also health-protective classification. Additionally, we validate our method by comparing the results against expert judgment-based classifications.

2. Methods

2.1. Simulation of maximum plasma concentrations of five example compounds

For the simulation of five example compounds, identical physiologically based kinetic (PBK) models using a generic hepatic clearance process and the PK-Sim Standard partitioning and permeability methods were generated in PK-Sim version 11.1.137 (Lippert et al., 2019; Willmann et al., 2003). For simplicity, it was assumed that all five example compounds have identical molecular weights of 300 g/mol, lipophilicity of 1 and fraction unbound values of 50 %. Solubility, intestinal permeability and clearance values of compounds differed as presented in Table 1. Then, daily oral administration of solutions of varying doses (1×10^{-4} to 1×10^3 mg/kg) were simulated for 5 years, and the peak plasma concentration for each dose was extracted. The oral bioavailability of the example compounds was predicted by use of the corresponding standard functionality of PK-Sim, relying on the generated PBK models.

2.2. HT-PBK modelling of EPAA NAM designation compounds

High-throughput PBK (HT-PBK) simulations of EPAA NAM Designation compounds were based on previously performed model validations (Gadaleta et al., 2024; Geci et al., 2024). Briefly, all compound properties required for PBK modelling (Kuepfer et al., 2016) were predicted using previously evaluated *in silico* tools: ADMETLab logD, logS and fraction unbound (Fu et al., 2024), ADMET Predictor v12 Fasted State Simulated Intestinal Fluid solubility and intrinsic liver clearance (<https://www.simulations-plus.com/software/admetpredictor>), OPERA intrinsic liver clearance (Mansouri et al., 2018). If available, we also retrieved for comparison *in vitro* measured fraction unbound and hepatocyte clearance values of compounds from htkk 2.4.0 (Pearce et al., 2017; Wetmore et al., 2012). For each compound, the default whole-body PBK model implemented in PK-Sim version 11.1.137 was parameterised from R version 4.2.2 (R Core Team, 2022), and administration of $1 \mu\text{g/kg}$ bodyweight once daily oral solutions were simulated, from which the total amount of compound in the body after five years was extracted as a readout. For partitioning, the method of Berezhevskiy (2004) was used. Peak Systemic Availability Factor (PSAF) and Mean Systemic Availability Factor (MSAF) values were calculated as shown in Eqs. (1) and (2).

Table 1

Properties of five example compounds of different Potential Systemic Availability concerns and different oral bioavailability. Other compound properties are assumed to be identical: molecular weight of 300 g/mol, lipophilicity of 1, fraction unbound of 50 %.

Compound No.	1	2	3	4	5
Potential Systemic Availability	Low	Medium (Low Bioavailability)	Medium (Intermediate Bioavailability)	Medium (High Bioavailability)	High
Clearance	15 ml/min/kg	15 ml/min/kg	7.5 ml/min/kg	0.5 ml/min/kg	1×10^{-5} ml/min/kg
Solubility (logS) [log 10 mol/l]	-5	-2	-2	-2	-2
Intestinal Permeability [log 10 cm/s]	-9	-6	-6	-6	-6
Predicted Oral Bioavailability (F)	~0%	~5%	~53%	~97%	~100%

2.3. Manual annotation of EPAA NAM designation compounds

For all 150 EPAA NAM Designation compounds, we searched for publicly available data that could be informative of their toxicokinetic properties, including *in vitro* data, *in vivo* data from animal and human studies, as well as potentially relevant regulatory classifications such as persistent organic pollutant (POP), persistent bioaccumulative toxic (PBT), very persistent very bioaccumulative (vPvB), and substance of very high concern (SVHC). Further, we extracted basic molecular properties like molecular weight, predicted water solubility and logP values of compounds using above mentioned *in silico* tools.

After having collected potentially relevant data, we decided to classify all endogenous compounds, such as sugars and amino acids, as medium PSA concern. Next, all compounds which possessed TK data and did not show signs of meaningful accumulation were also classified as medium concern. Compounds which did show signs of meaningful accumulation over repeated exposures were classified as high PSA concern. Due to highly heterogeneous TK data, e.g., variations in species, administration protocols, investigated tissues, defining absolute criteria for “meaningful” accumulation was not feasible and instead had to be determined by experts on a case-by-case-basis holistically considering all available information. Then, in the absence of more information, all remaining compounds without TK data that fell into the “small molecule” space were also classified as medium PSA concern. The two compounds which fell out of this space, both due to large molecular weights above 1500 g/mol, were classified as low concern.

3. Results

3.1. Definition and rationale of potential systemic availability concern classes

“Potential Systemic Availability” (PSA) is considered an important factor for the assessment of chemical safety in the framework suggested by Berggren and Worth (2023). While the reasoning behind its consideration is clear, the term itself remained underdefined. It was stated that it should encompass toxicokinetics, ADME processes, bioaccumulation, and persistence. However, no official definition exists, nor examples of chemicals representing the different concern categories. To address this gap, we first propose a definition of PSA and justify our definition in direct contrast with a potential alternative.

The definition of PSA concerns we will use in the following is that compounds shall be considered...

... “Low PSA Concern” if their *in vivo* concentrations can be assumed to never reach levels required for bioactivity under almost all relevant exposure scenarios, so that such compounds can be assumed to be safe under almost all conditions.

... “Medium PSA Concern” if the reaching of critical *in vivo* concentrations depends on the specific exposure scenario and the bioactivity of the compound. For such compounds, we assume that there will be an exposure dependency for their toxicity, with the compound being safe below a certain exposure threshold but unsafe above it.

... “High PSA Concern” if already very small exposures can lead to *in vivo* concentrations required for bioactivity, even if the bioactivity of the compound is only modest. Such compounds would pose a risk under almost all relevant exposure scenarios.

This definition may be understood as a “hard” interpretation of what property the PSA concerns should reflect, as it sets a very high standard for compounds to be classified as either low or high concern. By this definition, most small molecule chemicals, including almost all drugs, pesticides, etc., would fall into the medium PSA concern class, and only extreme chemicals would be considered low or high concern. Further, our definition explicitly refers to peak concentrations *reaching* levels required for *in vivo* bioactivity, which may be transitory. It does not require that these concentrations are also maintained for prolonged periods. In other words, our definition does not rely on exposure over time, as in the area under the curve (AUC), or average concentrations of compounds, but their peak concentration values (C_{max}), making this a conservative definition of the term.

Next, we outline a “softer” interpretation of the term to contrast our definition with a potential alternative. This would be to equate PSA with the bioavailability of a compound, which in pharmacology is defined as the fraction of a compound that reaches systemic circulation. In this reading, one may suggest that compounds could be classified as low, medium, or high PSA concern if they have low, medium, or high bioavailability. Whatever the specific thresholds for low, medium and high classification are, it is evident that such an interpretation would lead to a more sensitive division of typical small molecules. Drugs, pesticides, etc., can have drastically varying levels of bioavailability, ranging from just a few percent to full (100 %) bioavailability. Hence,

when using such a “softer” definition, these compounds would be allocated more evenly into the different PSA concern classes than when using our “hard” interpretation of the term.

To justify why we will continue our work using the “hard” definition, we illustrate *in vivo* plasma concentrations of five example compounds (Table 1). For simplicity and ease of comparability, we assume the five compounds have mostly identical properties, like identical molecular weight, lipophilicity, etc., and only differ in three key parameters: their i) clearance, ii) solubility and iii) intestinal permeability. Using a simple generic PBK model, we then simulate the maximum plasma concentration of these compounds after repeated oral exposure to different once-daily doses for five years. Further, we add two horizontal lines representing thresholds for *in vivo* concentrations of high (10 nM) and low (100 µM) bioactivity for contextualisation. The simulation results (Fig. 2) show that, despite their large differences in bioavailability, compounds 2–4 can all reach both the low and high concentration thresholds required for bioactivity. Whether the compounds would do so *in vivo* solely depends on the specific exposure dose and the specific bioactivity of the compounds. Hence, we conclude that these compounds cannot *a priori* be assumed to be safe or unsafe, at least not just based on their PSA, as a “softer” definition would suggest.

However, the situation is different for compounds 1 and 5. Due to its oral absorption being permeability- and solubility-limited, compound 1 would indeed never be able to even reach the 10 nM (“very bioactive”) concentration threshold, regardless of what its daily oral exposure is. Hence, we consider it to be justified to *a priori* classify such a compound as being of generally lower concern. Compound 5 would possess equal *in vivo* concentrations as compound 4 after single oral exposure (SI-Fig. 1),

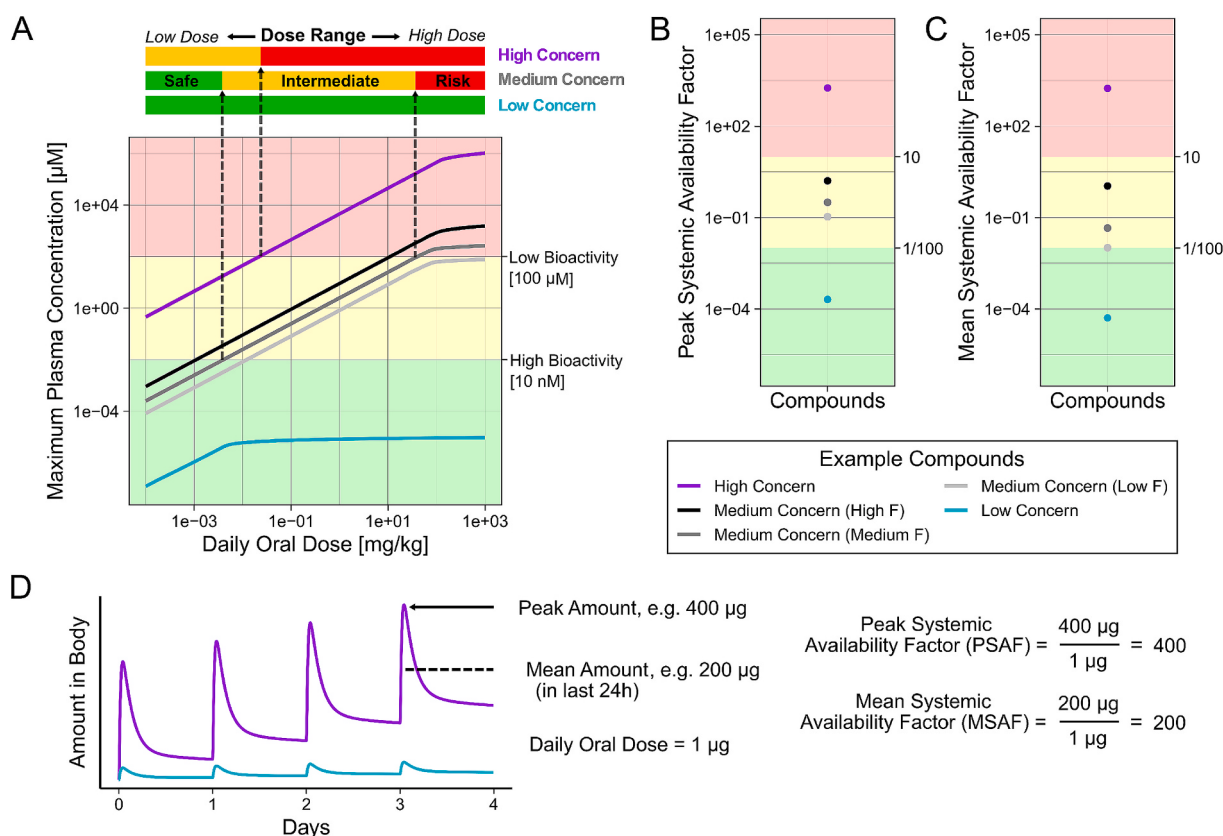


Fig. 2. Visual representation for the rationale behind distinguishing different Potential Systemic Availability concern classes. (A) Simulated maximum plasma concentration (C_{max}) against oral doses after daily administration for five years of five example compounds as described in Table 1. Marked as safe is the dose range for which the corresponding C_{max} value remains below the high bioactivity concentration (10 nM). Risk is the dose range where the C_{max} exceeds the low bioactivity threshold (100 µM), and intermediate is the area in between. (B) and (C) show the calculated Peak Systemic Availability and Mean Systemic Availability Factors of the five example compounds, respectively. (D) Example of Peak and Mean Amounts derived from a time-course profile, and calculation of the corresponding PSAF and MSAF values.

but due to its ability to bioaccumulate over time, it would reach much higher concentrations after prolonged periods of repeated exposure, even if the daily doses themselves are low. In the case of compound 5, this means that already at relatively modest exposures of 0.1 mg/kg, it would reach *in vivo* concentrations above the 100 μ M concentration threshold required for bioactivity, even though it is not very bioactive. Due to the increased risk of such compounds to reach concentrations required to cause adverse effects *in vivo*, and the undesirability of the bioaccumulation potential itself, we consider it warranted to *a priori* classify such compounds as being of higher concern.

3.2. PBK model-based quantification of potential systemic availability

Having clarified the definition of the PSA concern classes, we have a conceptually more rigorous foundation for the assessment of chemicals. However, from a practical perspective, the question of how this property can be evaluated, ideally efficiently for large numbers of chemicals, remains open.

Three considerations seem particularly important:

- i) The method should be organ-agnostic, meaning it should not only rely on concentrations in a single organ. While for illustrative purposes Fig. 2A gives the simplified impression that compounds only have a single *in vivo* concentration, the reality is more complex. The human body is not one homogenous compartment, but it is made up of multiple organs and compound concentrations between those organs can vary substantially. Typically, one

$$\text{MSAF} = \frac{\text{Mean Amount of Compound in the Body in the last 24h after 5 years of Daily Oral Exposure}}{\text{Daily Exposure Dose}} \quad (2)$$

would not *a priori* know which organs are potential targets of toxicity and which organ's concentration would be relevant for safety considerations. Hence, selecting a single organ for classification purposes is not trivial. It may even be counterproductive. For example, many in the body accumulating chemicals are highly lipophilic and, therefore, exhibit higher concentrations in fatty tissues, and, conversely, lower concentrations in plasma. If one only tracks concentrations in plasma, for example, then a more lipophilic, and for this reason bioaccumulating compound, may falsely appear less concerning than a less lipophilic compound.

- ii) To correctly capture the drastic differences in compounds' potential to bioaccumulate, it is required to consider repeated exposures over longer time periods, not just short-term or even single exposures.
- iii) The output should reflect a general, intrinsic compound property that is ideally independent of arbitrary assumptions or threshold values, e.g. specific exposure doses or concentration thresholds. Such arbitrary, or at least subjective, values can be controversial and may undermine the general acceptance of assessments.

Based on the outlined considerations, we define the Peak Systemic Availability Factor (PSAF) as

$$\text{PSAF} = \frac{\text{Peak Amount of Compound in the Body during 5 years of Daily Oral Exposure}}{\text{Daily Exposure Dose}} \quad (1)$$

This metric can be understood as the peak total body burden of a compound relative to its daily exposure, and it satisfies all three outlined requirements: (i) It considers the total amount of compound in the body, not just in arbitrarily selected organs. (ii) It considers accumulation after repeated exposure. (iii) The value itself is unitless and reflects an intrinsic compound property. Fig. 2B shows that the PSAF quantitatively captures the expected differences between the five example compounds: Compound 1 has the lowest PSAF of 2×10^{-4} , compounds 2 – 4 fall between 0.1 and 1.6, and compound 5 has a PSAF of 1789. Fig. 2D gives a visual representation of how this parameter would be derived from a specific time-course profile.

Despite avoiding arbitrary assumptions for the calculation of the PSAF itself, its translation from numerical values into the three PSA concern classes does eventually necessitate the definition of thresholds, which, as outlined before, are arbitrary. We here decide on a value of 10 for the distinction between medium/high concern and 1/100 for the distinction between low/medium. Practically, this means that a compound is considered of high concern if it can reach 10 times higher amounts in the body than the daily oral exposure to that compound is. Conversely, it would be of low concern if the total amount of compound in the body is never higher than 1 % of the daily oral exposure.

To explore and contrast the PSAF for scientific purposes with a second potential evaluation metric, we further define the Mean Systemic Availability Factor (MSAF) as

The MSAF is very similar to the PSAF, only that it does not evaluate the highest (peak) amounts of compound in the body, but its time-averaged mean amount. Similar to bioavailability, this makes the MSAF more sensitive to distinguishing chemicals than the PSAF. Compounds that are being cleared more rapidly can potentially have similar PSAF values as compounds that are cleared only slowly, but their MSAF values would be lower. For ease of comparison, we keep the same classification thresholds of 10 and 1/100, although, there may be good reasons to use different thresholds for the two metrics. Fig. 2C shows that the five example compounds can also be classified correctly by use of the MSAF. Although, the MSAF values of the three medium-concern compounds lay closer to the low-concern compound than it is the case for the PSAF values.

3.3. Plausibility evaluation of potential systemic availability classifications

To demonstrate the utility of our classification metrics, we calculate PSAF and MSAF values for a large set of diverse substances, the 150 EPAA NAM Designation compounds. For this, we use a previously described high-throughput PBK modelling approach (Geci et al., 2024), which enables the rapid generation of generic compound PBK models. This allows the high-throughput prediction of *in vivo* concentrations, and hence the efficient evaluation of our proposed scores for a large dataset.

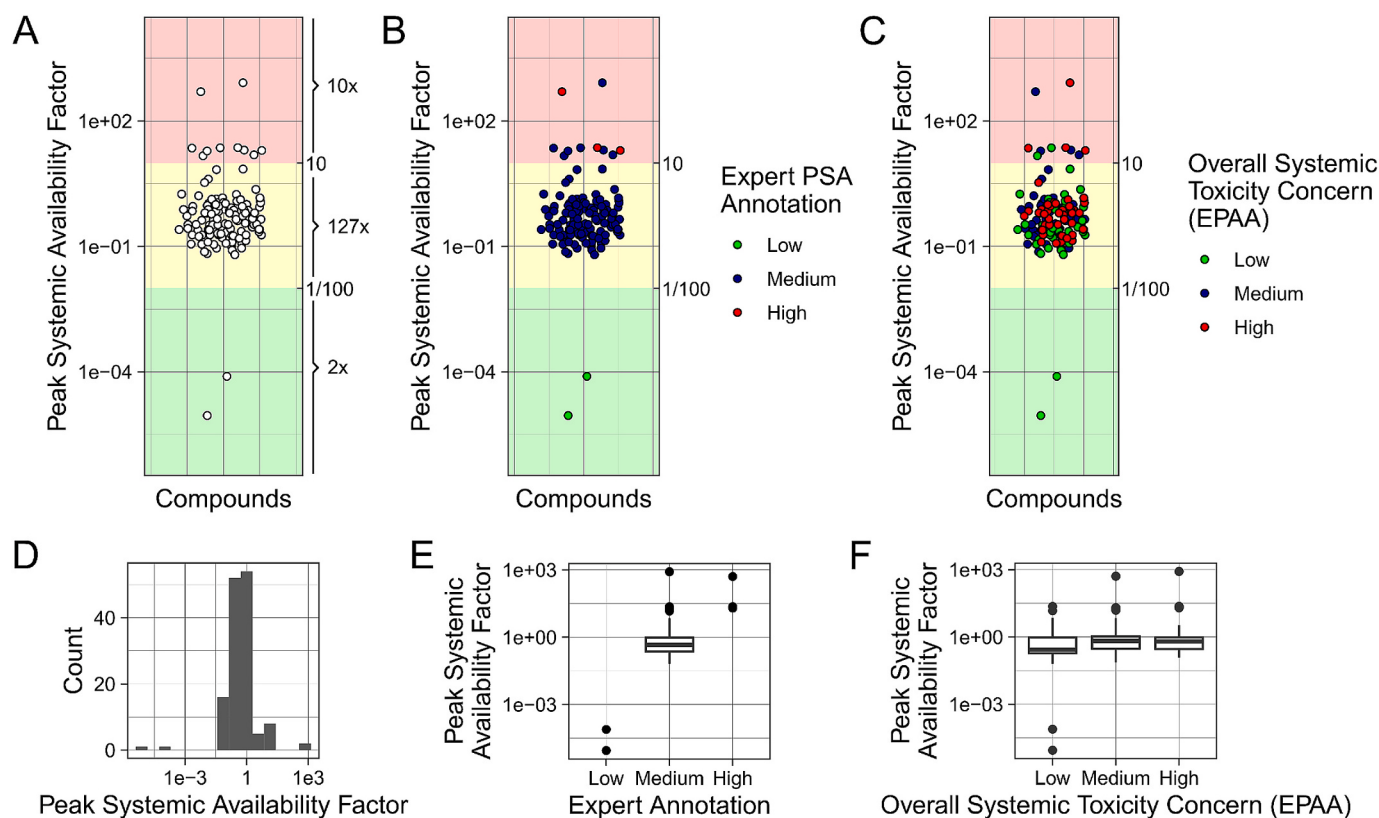


Fig. 3. Peak Systemic Availability Factors of the 139 EPAA Designation Compounds for which HT-PBK predictions could be made. (A,D) Distribution of PSAF values with jittered x-axis. (B) Coloured by manually annotated reference classification. (C) Coloured by overall systemic toxicity concern given by EPAA. (E, F) Comparison of PSAF values of the different manually annotated Potential Systemic Availability or overall concern classes.

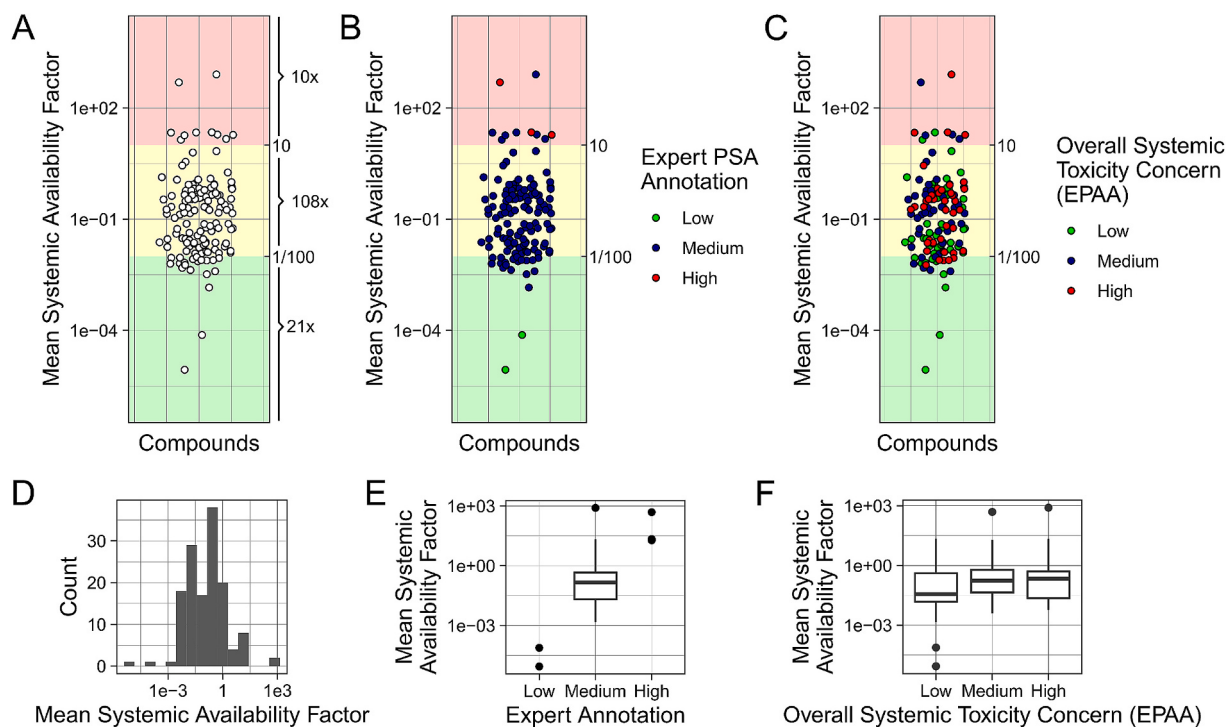


Fig. 4. Mean Systemic Availability Factors of the 139 EPAA Designation Compounds for which HT-PBK predictions could be made. (A,D) Distribution of MSAF values with jittered x-axis. (B) Coloured by manually annotated reference classification. (C) Coloured by overall systemic toxicity compound concern given by EPAA. (E, F) Comparison of MSAF values of the different manually annotated Potential Systemic Availability or overall concern classes.

Nonetheless, we would like to highlight that our method can be used with any type of PBK model and is independent of how the PBK models used are parameterised.

HT-PBK modelling-based predictions could be made for 139 out of the 150 EPAA NAM Designation compounds. The remaining 11 compounds were inorganics, metals or organometallics, for which the input *in silico* tools could not provide the required property predictions. Predicted PSAF and MSAF values of the EPAA NAM Designation compounds are shown in Fig. 3A and Fig. 4A, respectively. Only 2 out of the 139 successfully predicted compounds have a PSAF < 1 %, and hence are classified as low PSA concern. 10 compounds were predicted to have a PSAF > 10 and end up in the high concern category. The remaining 127 compounds all have PSAF values between 1 % and 10 and therefore are classified medium PSA concern. Using MSAF generally leads to similar classification results as PSAF, only that more compounds are considered low PSA concern (21 instead of 2). The number of high PSA concern compounds is identical to the PSAF. As expected, we observe that MSAF values are distributed more widely than PSAF values.

To have a comparison for assessing the plausibility of our predicted classifications, we manually established a reference classification based on review of public data and expert judgment of the NAM Designation compounds. The details of this process are outlined in Methods 2.3, and the manually annotated reference classifications are provided in [Supplementary File 2](#). It proved challenging to annotate compounds since the type, quality and consistency of data available for them were highly variable. For some compounds, high-quality *in vivo* data after repeated oral exposure in humans was available, while others only had lower-quality data from animal or *in vitro* studies, or not even any TK-relevant information at all.

We classified all endogenous compounds (amino acids, sugars, fatty acids) as medium PSA concern, assuming such endogenous substances would be absorbed rather efficiently, yet also not bioaccumulate. One potential exception to this was Ergocalciferol (Vitamin D2), for which a classification as high PSA concern was considered because it is a fat-soluble vitamin which can accumulate in the body over time. But since its half-life is only on the order of days, we eventually decided that its accumulation potential was still not high enough and hence also classified it as only medium PSA concern.

All other compounds with TK data available had been reported to be absorbed orally to some extent, at least so much that for none of them classification as low PSA concern seemed justified. Their reported clearances, half-lives or potential to accumulate in tissues, however, varied quite substantially. Compounds that had been reported to bioaccumulate strongly or have half-lives of at least weeks were then investigated more closely for potential classification as high PSA concern. We eventually classified three compounds as high PSA concern (Perfluorooctanoic acid, Perfluoroheptanoic acid, Hexabromocyclododecane), and the remaining ones as medium. We observed that two compounds (Polyoxyl(40)stearate and 2-Hydroxypropyl-beta-cyclodextrin) possessed very large molecular weight (>1500 g/mol) and classified both as low concern, assuming they are not absorbed orally to any relevant extent. All remaining compounds for

which we could not find specific TK data fell into the small molecule domain with less than 1000 g/mol molecular weight. In the absence of further evidence, we hence assumed these are all medium PSA concern. For some compounds, there was additional uncertainty in their correct reference classification due to special properties like depuration by evaporation (Decamethylcyclopentasiloxane; Gobas et al., 2015a) or potential explosiveness (Dipicrylamine). Eventually, we classified 2 of the compounds as low, 134 as medium, and 3 as high PSA concern.

Comparing our HT-PBK predictions of PSAF values to the expert judgement-based reference classifications revealed that the bulk of compounds had been classified correctly, with 132 out of 139 chemicals grouped identically (Table 2). 7 compounds that were manually classified as only medium PSA concern, were predicted to be of high concern by HT-PBK predictions. However, this type of false-positive misprediction is rather unproblematic for assessment purposes. No compound was predicted to have lower classification than decided by expert judgment. For MSAF-based classifications, however, we did find some compounds being falsely predicted to have low classification, while experts classified them medium PSA concern (SI-Table 2). Even after investigating these compounds in more depth, we did not find any evidence suggesting that these might be of low concern.

Eventually, we also compared predicted PSA classifications to the overall systemic toxicity concern provided by the EPAA. The overall concern categorisation considers not only the PSA but also the bioactivity of a substance and was derived using traditional toxicity data. Hence, these classifications are not expected to be identical to our PSA classifications. But since PSA is one of the two relevant factors of the overall systemic toxicity concern, the comparison may still give a rough indication of the plausibility of our PSA classifications. At least, there should be no compound of overall high concern that is only classified low PSA, and no compound that is of overall low concern but classified high PSA. Indeed, we observed that the first is not the case. The 2 compounds classified as low PSA are also classified low overall concern by the EPAA. However, out of the 10 compounds predicted to be high PSA by use of HT-PBK modelling and the PSAF, 2 compounds are classified only low overall concern. Assuming the overall concern classification by the EPAA is correct, this suggests that the HT-PBK based classification of these compounds is too high, and that these are false positives.

The two false positive compounds were Decamethylcyclopentasiloxane (CAS 541-02-6) and Sodium myristoyl glutamate (CAS 38517-37-2). Both showed high predicted PSAF values because their hepatic clearance was predicted to be zero by OPERA (Mansouri et al., 2018). For Decamethylcyclopentasiloxane, this prediction is consistent with the *in vitro* hepatic clearance value reported in htk (Pearce et al., 2017), and it has also been extensively investigated for its potential to bioaccumulate (Gobas et al., 2015b; Kierkegaard et al., 2013). However, studies have shown that in air-breathing organisms, it is eliminated through exhalation (Gobas et al., 2015a). This elimination pathway is not represented in our generic HT-PBK models. As a result, the excretion of Decamethylcyclopentasiloxane was likely underpredicted. For Sodium myristoyl glutamate, no *in vitro* hepatic clearance data are available from htk. However, we assume that a similar situation may apply.

Generally, we observed that there is a tendency of low overall concern compounds to have lower PSAF and MSAF values. This effect is more pronounced for MSAF values, although there is still large overlap between the three overall concern groups. We also observed that, unlike for the PSAF, some high overall concern compounds have only low MSAF values. Hence, it seems that the more strongly spread distribution of MSAF values, and its potential ability to better distinguish lower concern compounds comes at the risk of causing more false-negative classifications. Only the more conservative peak-based classification resulted in no false-negative classifications.

Table 2

Comparison of classifications of 139 EPAA Designation Compounds based on HT-PBK predictions using the Peak Systemic Availability Factor (PSAF) against expert judged classification of Potential Systemic Availability (PSA).

	HT-PBK PSAF: Low	HT-PBK PSAF: Medium	HT-PBK PSAF: High	Sum
Expert PSA class: Low	2	0	0	2
Expert PSA class: Medium	0	127	7	134
Expert PSA class: High	0	0	3	3
Sum	2	127	10	139

4. Discussion

We here proposed a definition of PSA concern classes, and a method for quantitatively measuring this property. Based on expert judgement, we manually annotated a benchmark dataset of diverse chemicals from the EPAA NAM Designathon and compared our predicted classifications to this annotation. This demonstrated that HT-PBK modelling could be used to correctly predict chemical PSA classifications in a high-throughput manner.

The conceptual rationale for our classification of compounds directly builds on the fundamental dogma of toxicology that chemicals can only cause adverse effects if their doses, and hence their *in vivo* concentrations, are sufficiently high (“the dose makes the poison”). We show that certain compounds can be assumed to be unlikely to reach such *in vivo* concentrations, and we argue that therefore they deserve to be treated more permissively in chemical safety assessment. Conversely, we showed that compounds that have the potential to accumulate in the body over time can reach harmful *in vivo* concentrations already at low daily doses, even if they only possess modest bioactivity. Hence, those deserve to be treated more cautiously.

Further, the potential for bioaccumulation comes with additional implications that are undesirable from a safety perspective. For instance, if the bioactivity of a bioaccumulating substance is underestimated, there may be a significant delay between the exposure to chemicals and the observation of their toxic effects, as it may take years for substances to accumulate to steady-state levels in humans. Subsequently, their adverse effects may only become apparent long after widespread exposure. By this time, these chemicals may have already been used and widely distributed, making retroactive mitigation and response efforts highly challenging. This is currently a major problem regarding per- and polyfluoroalkyl substances (PFAS; Brunn et al., 2023; Sunderland et al., 2019).

Our definition of PSA reflects a health-protective interpretation of the term as it assumes that reaching peak levels once may be sufficient to trigger adverse effects *in vivo*. However, evidence suggests that for some chemicals, toxicity is more closely linked to exposure over time (AUC), rather than peak (C_{max}) levels (Groothuis et al., 2015; Tzimas et al., 1997). This raises the possibility that AUC may be a better metric for the classification of chemicals. To explore this possibility, we investigated whether a time-average-based approach is useful for classification purposes. Our findings demonstrate that the most straightforward approach using bioavailability as a proxy for PSA, is not suitable for classification. Instead, we proposed a method that accounts for exposure across all organs and considers accumulation from repeated exposures over time.

For our benchmark chemical set, we find that peak (PSAF) and time-averaged (MSAF) classifications yield comparable results, with most NAM Designathon compounds clustering around moderate values with only a few extremes. Hence, both metrics could potentially be used for the classification of chemicals. Further, we observe that compounds of low overall systemic toxicity concern tend to also have lower MSAF values, and the distinction between those and higher overall concern compounds is more pronounced with MSAF than with the PSAF metric. This seems reasonable, as compounds that exhibit lower exposure have, all else being equal, less potential for causing adverse effects.

However, we also find exceptions to this trend, where high overall concern compounds display low MSAF values. We speculate that those represent cases where chemical toxicity is driven by peak levels, rather than exposure over time. For instance, phosphamidon is an organophosphate pesticide, which are cleared relatively efficiently (Moerenhout et al., 2024), and MSAF-based classification also suggests it is of only low PSA concern. In the context of the EPAA Designathon, it is classified as overall high systemic toxicity concern. And indeed, phosphamidon is a known cholinesterase inhibitor that causes adverse effects above certain concentration thresholds (Sidhu et al., 2019). Therefore, it seems plausible that this compound’s toxicity is driven by C_{max}, rather than AUC, and that it is only classified appropriately by

consideration of peak levels, but not by time-averaged levels. Given that it is typically unknown beforehand whether peak concentration or exposure over time is the primary driver of toxicity for a given compound, we conclude that a peak-based approach is better suited for safety assessments, as it is the more conservative metric and avoids false-negative classifications, which may arise when using time-average-based values. Even though false-negative classifications may be rare, they can undermine public and regulatory acceptance.

Further, we manually annotated the PSA classes of the Designathon compounds to have a reference for the evaluation of our predicted classifications. Consistent with our predictions, we concluded that most compounds are best classified as medium PSA concern. These compounds, therefore, cannot *a priori* be considered more or less safe, since they may end up in any of the three systemic toxicity concern categories, depending on their specific bioactivity. Considering that most of the investigated compounds are small molecules, like endogenous metabolites, food ingredients, drugs, and pesticides, this seems plausible. These substances have been deliberately selected by nature or industry to be bioavailable and/or exert bioactive effects. For example, the food additive Tartaric acid, the drug Ketoconazole and the pesticide Glufosinate ammonium have similar PSAF values of ~80 % and are all classified as medium PSA concern, despite being classified differently regarding their overall concern by the EPAA. Tartaric acid is only of low overall concern, whereas Glufosinate ammonium and Ketoconazole were both classified as high overall concern. On the other hand, this observation may potentially be counterintuitive, as the NAM Designathon compounds were selected to equally represent the three overall systemic toxicity concern classes, with 50 compounds from each class. Hence, one might have expected them to be also distributed evenly across the three PSA concern classes.

It is important to note that the data available for establishing reference classifications was highly heterogeneous, making it challenging to confidently assign reference classes to many compounds. We acknowledge the subjectivity of our manual annotation process, which relied on interpreting heterogeneous data on a case-by-case-basis, without established standards or protocols. Nevertheless, we believe the general trends described here are robust. Alternative data sources, such as human biomonitoring data, may provide insights for more accurate classification of some chemicals, although such an analysis was beyond the scope of this study.

Because we evaluated PSAF and MSAF values by using HT-PBK modelling, our results are subject to the general limitations of this method. Primarily, the accuracy of our predictions depends on the reliability and applicability of the underlying *in silico* tools providing the input parameters for PBK models. When input parameters are incorrect, the simulation outputs will also be flawed. Addressing these limitations requires a sustained effort to further validate *in silico* tools, to expand their applicability domains, and to improve their reliability and accuracy. For the shown application, it is especially relevant that *in silico* tools not just accurately predict average values but also extreme outliers, such as compounds with exceptionally low or high clearance, fraction unbound, or intestinal permeability values. These extreme properties are often the primary drivers for classification into low or high PSA concern classes. Unfortunately, *in silico* tools currently struggle precisely with this, since their training datasets often underrepresent compounds having extreme values (Cherkasov et al., 2014; Kirchmair et al., 2015). Addressing this will be essential for improving the accuracy of *in silico* classifications like those presented here. Alternatively, our PBK models may be parameterised using *in vitro* data instead. However, *in vitro* methods also possess biases and may potentially face similar difficulties in handling extreme chemicals (Houston and Galetin, 2003; Proença et al., 2021). When we compared our *in silico* prediction-based PBK model simulations to simulations parameterised using *in vitro* measurements of fraction unbound and hepatocyte clearance, for a subset of compounds with available data from the US EPA’s httk package (Pearce et al., 2017; Wetmore et al., 2012), we observed no meaningful

differences (Supplementary Fig. 2).

Limitations of the *in silico* tools used here also explain why certain groups of substances, like metals or inorganics, could not be assessed with our approach. Nevertheless, the coverage of 139 out of 150 compounds of interest was high, suggesting this may not be typically problematic for applications in safety assessment. Additionally, we identified some non-standard processes that may influence compounds' PSA concern but were not included in our generic PBK models. Examples include the depuration of volatile compounds by breathing, chemical instability, gut metabolism, and biliary clearance. Ideally, all relevant processes should be included in PBK models to enable comprehensive and reliable evaluations. However, such detailed modelling was beyond the scope of our analysis and the development of predictive tools for these processes remains an ongoing challenge. The same applies to the prediction of transporter activities, which can significantly affect local organ concentrations and sometimes may be a crucial factor for *in vivo* toxicity (Klaassen and Aleksunes, 2010; Najjar et al., 2022).

Finally, a comprehensive assessment of total human exposure should encompass multiple exposure routes, including, for instance, dermal and inhalation exposure. Here, we focused solely on oral absorption, as it represents one highly relevant route of exposure. Nevertheless, our approach can be applied in the same way to other exposure routes. Generally, PBK modelling enables the comparison, or integration, of exposures of multiple routes and across different demographic groups to facilitate comprehensive assessments of chemical exposure (Li and Xiong, 2023).

An alternative to the PBK modelling-based approach is to use rule-based classification schemes or simple regression models relying on selected chemical properties. This is, for example, done for predicting bioaccumulation in aquatic organisms and bioconcentration factors (Arnot and Gobas, 2006). However, this requires the subjective definition of classification thresholds and the selection of individual compound properties to be considered (Arnot et al., 2010). In contrast, PBK modelling possess the power to integrate various relevant molecular properties that jointly determine quantitative *in vivo* behaviours. It is therefore better suited to capture quantitative *in vivo* outcomes and the complex relationships between the various molecular properties. This is why PBK modelling is increasingly applied to evaluate *in vivo* exposure to chemicals (Coecke et al., 2013; Tonnelier et al., 2012; Wambaugh et al., 2015; Wetmore et al., 2012), although ultimately the choice of method may depend on the context, available information and the required accuracy.

We observed for both predicted and manually annotated classifications that most compounds should be considered medium PSA concern, yet their overall systemic toxicity concerns are distributed equally among the three classes. This suggests that the overall concern of compounds is driven mainly by differences in compound bioactivity, which questions the utility of TK-based classifications. Regarding this, we would like to point out two things. Firstly, identifying high PSA concern compounds is valuable, even if they only represent a small subgroup of chemicals, because it enables prioritised assessments of such compounds. This is a significant benefit given the existence of thousands of untested chemicals on the market. Secondly, when considering the broader chemical universe, it is conceivable that a substantial number of chemicals may fall into the low PSA category. But, potentially, these chemicals were just not represented in the NAM Designathon dataset, and are not even routinely tested, because i) their extreme chemical properties, such as low permeability, make them difficult to test *in vitro*, and ii) because based on their low bioavailability they are not expected to cause relevant adverse effects that would meet existing classification criteria. Therefore, our classification approach may simply reflect the current handling of such compounds and supports the rationale for their current regulatory treatment. For example, under our definition of PSA, most polymers would likely be classified as low concern, as they are generally regarded to not readily cross biological membranes and hence to exhibit minimal systemic toxicity. This

explains why polymers are subject to regulatory exceptions (U.S. Environmental Protection Agency, 1997), aligning well with our results, yet such compounds were not represented in our evaluation dataset.

Finally, the application of the presented method remains to be explored more concretely within the context of a novel framework for regulatory decision-making. It is an ongoing effort to outline how the current system of chemical safety assessment may be changed to finally phase out animal testing (Ball et al., 2022; Berggren and Worth, 2023; Schmeisser et al., 2023; Worth et al., 2024; Worth and Berggren, 2024). We showed that the consideration of TK properties in the form of the PSAF could support an *a priori* prioritisation of certain chemicals, and the de-prioritisation of others. This may be useful for classifying, or at least triaging, compounds within a novel tiered safety assessment framework. Since the here used high-throughput PBK model approach only requires *in silico* predictions of a few compound properties, it allows the rapid classification of the large number of currently unassessed substances. Even if one would use more reliable *in vitro* data for parameterisation, it would still be possible to efficiently conduct such an analysis for large numbers of compounds (Rotroff et al., 2010; Wetmore et al., 2012).

However, we also found that many compounds cannot be (de)prioritised outright, as most were classified into the medium PSA concern category. From this category, classification into any overall systemic toxicity concern class remains possible, depending on the bioactivity of compounds. In cases where the bioactivity of compounds is also neither particularly low nor high, it would be necessary to conduct a more in-depth comparison of their *in vivo* concentrations after specific exposures against their bioactivities to determine bioactivity exposure ratios (BERs) or concrete thresholds for (un)safe exposure levels (Wetmore et al., 2015). Such more elaborate, in-depth evaluations could build directly on the same HT-PBK models used for an initial PSA classification.

In the absence of information on specific bioactivity and internal exposure levels, the presented approach had to be based on highly conservative assumptions. We anticipate that integrating bioactivity data into the analysis will eventually enable a larger number of NAM Designathon compounds to be classified more definitively as either low or high systemic toxicity concern (Doe et al., 2025). This will finally enable a comparison of overall systemic toxicity concern classifications using NAMs against classifications based on traditional toxicological studies.

CRedit authorship contribution statement

René Geci: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alicia Paini:** Methodology, Investigation, Data curation, Conceptualization, Writing – review & editing, Supervision. **Andrew P Worth:** Writing – review & editing, Methodology, Investigation, Data curation. **Lars Kuepfer:** Writing – review & editing, Visualization, Supervision. **Stephan Schaller:** Writing – review & editing, Supervision, Funding acquisition.

Funding

This work was performed in the context of the ONTOX project (<https://ontoxproject.eu/>) that has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 963845. ONTOX is part of the ASPIS project cluster (<https://aspiscluster.eu/>).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rene Geci reports financial support was provided by European Union.

Stephan Schaller reports financial support was provided by European Union. Rene Geci reports a relationship with University Hospital RWTH Aachen that includes: employment. Alicia Pains reports a relationship with ESQlabs GmbH that includes: employment. Andrew P Worth reports a relationship with European Commission Joint Research Centre Ispra that includes: employment. Lars Kuepfer reports a relationship with University Hospital RWTH Aachen that includes: employment. Stephan Schaller reports a relationship with ESQlabs GmbH that includes: employment and equity or stocks. Rene Geci reports a relationship with ESQlabs GmbH that includes: non-financial support and travel reimbursement. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2025.109636>.

Data availability

All data is available in the [Supplementary Information](#).

References

- Arnot, J.A., Arnot, M.I., Mackay, D., Couillard, Y., MacDonald, D., Bonnell, M., Doyle, P., 2010. Molecular size cutoff criteria for screening bioaccumulation potential: fact or fiction? *Integr. Environ. Assess. Manag.* 6, 210–224. <https://doi.org/10.1897/IEAM.2009-051.1>.
- Arnot, J.A., Gobas, F.A., 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. *Environ. Rev.* 14, 257–297. <https://doi.org/10.1139/a06-005>.
- Ball, N., Bars, R., Botham, P.A., Cuciureanu, A., Cronin, M.T.D., Doe, J.E., Dudzina, T., Gant, T.W., Leist, M., van Ravenzwaay, B., 2022. A framework for chemical safety assessment incorporating new approach methodologies within REACH. *Arch. Toxicol.* 96, 743–766. <https://doi.org/10.1007/s00204-021-03215-9>.
- Berezhkovskiy, L.M., 2004. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J. Pharm. Sci.* 93, 1628–1640. <https://doi.org/10.1002/jps.20073>.
- Berggren, E., Worth, A.P., 2023. Towards a future regulatory framework for chemicals in the European Union - chemicals 2.0. *Regulatory toxicology and pharmacology. RTP* 142, 105431. <https://doi.org/10.1016/j.yrtph.2023.105431>.
- Botham, P., Cronin, M.T.D., Currie, R., Doe, J., Funk-Weyer, D., Gant, T.W., Leist, M., Marty, S., van Ravenzwaay, B., Westmoreland, C., 2023. Analysis of health concerns not addressed by REACH for low tonnage chemicals and opportunities for new approach methodology. *Arch. Toxicol.* 97, 3075–3083. <https://doi.org/10.1007/s00204-023-03601-5>.
- Brunn, H., Arnold, G., Körner, W., Rippen, G., Steinhäuser, K.G., Valentin, I., 2023. PFAS: forever chemicals—persistent, bioaccumulative and mobile. Reviewing the status and the need for their phase out and remediation of contaminated sites. *Environ. Sci. Eur.* 35, 1–50. <https://doi.org/10.1186/s12302-023-00721-8>.
- Cherkasov, A., Muratov, E.N., Fourches, D., Varnek, A., Baskin, I.I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y.C., Todeschini, R., Consonni, V., Kuz'min, V.E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A., Tropsha, A., 2014. QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.* 57, 4977–5010. <https://doi.org/10.1021/jm4004285>.
- Coecke, S., Pelkonen, O., Leite, S.B., Bernauer, U., Bessems, J.G., Bois, F.Y., Gundert-Remy, U., Loizou, G., Testai, E., Zaldívar, J.-M., 2013. Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicol. Vitro: Int. J. Publ. Assoc. BIBRA* 27, 1570–1577. <https://doi.org/10.1016/j.tiv.2012.06.012>.
- Doe, J.E., Botham, P., Holland, D., Gatnik, M.F., Giri, V., Kang, H., Kalra, P., León Pérez, S., Marty, S., Moors, S., Raeburn, R., Reale, E., Settivari, R., Sica, M., Travis, K. Z., Wijeyesakere, S.J., 2025. Framework for classifying chemicals for repeat dose toxicity using NAMs. *Arch. Toxicol.* 1–24. <https://doi.org/10.1007/s00204-025-04069-1>.
- EEA, 2019. The European environment - state and outlook 2020: knowledge for transition to a sustainable Europe.
- EU, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH).
- Fu, L., Shi, S., Yi, J., Wang, N., He, Y., Wu, Z., Peng, J., Deng, Y., Wang, W., Wu, C., Lyu, A., Zeng, X., Zhao, W., Hou, T., Cao, D., 2024. ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic Acids Res.* 52, W422–W431. <https://doi.org/10.1093/nar/gkac236>.
- Gadaleta, D., Serrano-Candelas, E., Ortega-Vallbona, R., Colombo, E., Garcia de Lomana, M., Biava, G., Aparicio-Sánchez, P., Roncaglioni, A., Gozalbes, R., Benfenati, E., 2024. Comprehensive benchmarking of computational tools for predicting toxicokinetic and physicochemical properties of chemicals. *J. Cheminform* 16, 145. <https://doi.org/10.1186/s13321-024-00931-z>.
- Geci, R., Gadaleta, D., de Lomana, M.G., Ortega-Vallbona, R., Colombo, E., Serrano-Candelas, E., Pains, A., Kuepfer, L., Schaller, S., 2024. Systematic evaluation of high-throughput PBK modelling strategies for the prediction of intravenous and oral pharmacokinetics in humans. *Arch. Toxicol.* 98, 2659–2676. <https://doi.org/10.1007/s00204-024-03764-9>.
- Gobas, F.A.P.C., Powell, D.E., Woodburn, K.B., Springer, T., Huggett, D.B., 2015a. Bioaccumulation of decamethylpentacyclosiloxane (D5): a review. *Environ. Toxicol. Chem.* 34, 2703–2714. <https://doi.org/10.1002/etc.3242>.
- Gobas, F.A.P.C., Xu, S., Kozerski, G., Powell, D.E., Woodburn, K.B., Mackay, D., Fairbrother, A., 2015b. Fugacity and activity analysis of the bioaccumulation and environmental risks of decamethylcyclopentasiloxane (D5). *Environ. Toxicol. Chem.* 34, 2723–2731. <https://doi.org/10.1002/etc.2942>.
- Groothuis, F.A., Heringa, M.B., Nicol, B., Hermens, J.L.M., Blaauwboer, B.J., Kramer, N.I., 2015. Dose metric considerations in vitro assays to improve quantitative in vitro-in vivo dose extrapolations. *Toxicology* 332, 30–40. <https://doi.org/10.1016/j.tox.2013.08.012>.
- Houston, J.B., Galetin, A., 2003. Progress towards prediction of human pharmacokinetic parameters from in vitro technologies. *Drug Metab. Rev.* 35, 393–415. <https://doi.org/10.1081/dmr-120026870>.
- Kierkegaard, A., Bignert, A., McLachlan, M.S., 2013. Bioaccumulation of decamethylcyclopentasiloxane in perch in Swedish lakes. *Chemosphere* 93, 789–793. <https://doi.org/10.1016/j.chemosphere.2012.10.050>.
- Kirchmair, J., Göller, A.H., Lang, D., Kunze, J., Testa, B., Wilson, I.D., Glen, R.C., Schneider, G., 2015. Predicting drug metabolism: experiment and/or computation? *Nature reviews. Drug Discov.* 14, 387–404. <https://doi.org/10.1038/nrd4581>.
- Klaassen, C.D., Aleksunes, L.M., 2010. Xenobiotic, bile acid, and cholesterol transporters: function and regulation. *Pharmacol. Rev.* 62, 1–96. <https://doi.org/10.1124/pr.109.002014>.
- Kuepfer, L., Niederal, C., Wendt, T., Schlender, J.-F., Willmann, S., Lippert, J., Block, M., Eissing, T., Teutonico, D., 2016. Applied concepts in PBPK modeling: how to build a PBPK/PD model. *CPT Pharmacometrics Syst. Pharmacol.* 5, 516–531. <https://doi.org/10.1002/psp4.12134>.
- Li, Z., Xiong, J., 2023. A dynamic inventory database for assessing age-, gender-, and route-specific chronic internal exposure to chemicals in support of human exposure research. *J. Environ. Manage.* 339, 117867. <https://doi.org/10.1016/j.jenvman.2023.117867>.
- Lippert, J., Burghaus, R., Edgington, A., Frechen, S., Karlsson, M., Kovar, A., Lehr, T., Milligan, P., Nock, V., Ramusovic, S., Riggs, M., Schaller, S., Schlender, J., Schmidt, S., Sevestre, M., Sjögren, E., Solodenko, J., Staab, A., Teutonico, D., 2019. Open systems pharmacology community—an open access, open source, open science approach to modeling and simulation in pharmaceutical sciences. *CPT Pharmacometrics Syst. Pharmacol.* 8, 878–882. <https://doi.org/10.1002/psp4.12473>.
- Mansouri, K., Grulke, C.M., Judson, R.S., Williams, A.J., 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. *J. Cheminf.* 10, 10. <https://doi.org/10.1186/s13321-018-0263-1>.
- Moerenhout, T.M.J.A., Chen, J., Bouwmeester, H., Rietjens, I.M.C.M., Kramer, N.I., 2024. Development of a generic physiologically based kinetic model for the prediction of internal exposure to organophosphate pesticides. *Environ. Sci. Tech.* <https://doi.org/10.1021/acs.est.4c06534>.
- Najjar, A., Punt, A., Wambaugh, J., Pains, A., Ellison, C., Fragki, S., Bianchi, E., Zhang, F., Westerhout, J., Mueller, D., Li, H., Shi, Q., Gant, T.W., Botham, P., Bars, R., Piersma, A., van Ravenzwaay, B., Kramer, N.I., 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.* <https://doi.org/10.1007/s00204-022-03356-5>.
- Pearce, R.G., Setzer, R.W., Strobe, C.L., Wambaugh, J.F., Sipes, N.S., 2017. htk: R #ICS. *J. Stat. Softw.* 79, 1–26. <https://doi.org/10.18637/jss.v079.i04>.
- Pronça, S., Escher, B.I., Fischer, F.C., Fisher, C., Grégoire, S., Hewitt, N.J., Nicol, B., Pains, A., Kramer, N.I., 2021. Effective exposure of chemicals in in vitro cell systems: a review of chemical distribution models. *Toxicol. Vitro: Int. J. Publ. Assoc. BIBRA* 73, 105133. <https://doi.org/10.1016/j.tiv.2021.105133>.
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Rotroff, D.M., Wetmore, B.A., Dix, D.J., Ferguson, S.S., Clewell, H.J., Houck, K.A., Lecluyse, E.L., Andersen, M.E., Judson, R.S., Smith, C.M., Sochaski, M.A., Kavlock, R. J., Boellmann, F., Martin, M.T., Reif, D.M., Wambaugh, J.F., Thomas, R.S., 2010. Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol. Sci.: Off. J. Soc. Toxicol.* 117, 348–358. <https://doi.org/10.1093/toxsci/kfq220>.
- Schmeisser, S., Miccoli, A., von Bergen, M., Berggren, E., Braeuning, A., Busch, W., Desaintes, C., Gourmelon, A., Grafström, R., Harrill, J., Hartung, T., Herzler, M., Kass, G.E.N., Kleinstreuer, N., Leist, M., Luijten, M., Marx-Stoelting, P., Poetz, O., van Ravenzwaay, B., Roggeband, R., Rogiers, V., Roth, A., Sanders, P., Thomas, R.S., Marie Vinggaard, A., Vinken, M., van de Water, B., Luch, A., Tralau, T., 2023. New approach methodologies in human regulatory toxicology - not if, but how and when! *Environ. Int.* 178, 108082. <https://doi.org/10.1016/j.envint.2023.108082>.
- Sidhu, G.K., Singh, S., Kumar, V., Dhanjal, D.S., Datta, S., Singh, J., 2019. Toxicity, monitoring and biodegradation of organophosphate pesticides: a review. *Crit. Rev. Environ. Sci. Technol.* 49, 1135–1187. <https://doi.org/10.1080/10643389.2019.1565554>.
- Sunderland, E.M., Hu, X.C., Dassuncao, C., Tokranov, A.K., Wagner, C.C., Allen, J.G., 2019. A review of the pathways of human exposure to poly- and perfluoroalkyl

- substances (PFASs) and present understanding of health effects. *J. Exposure Sci. Environ. Epidemiol.* 29, 131–147. <https://doi.org/10.1038/s41370-018-0094-1>.
- Tonneller, A., Coecke, S., Zaldivar, J.-M., 2012. Screening of chemicals for human bioaccumulative potential with a physiologically based toxicokinetic model. *Arch. Toxicol.* 86, 393–403. <https://doi.org/10.1007/s00204-011-0768-0>.
- Tzimas, G., Thiel, R., Chahoud, I., Nau, H., 1997. The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol. Appl. Pharmacol.* 143, 436–444. <https://doi.org/10.1006/taap.1997.8105>.
- U.S. Environmental Protection Agency, 1997. Polymer Exemption Guidance Manual for Premanufacture Notices for New Chemicals under the Toxic Substances Control Act (TSCA): A Technical Manual to Accompany, but not Supersede the “Premanufacture Notification Exemptions; Revisions of Exemptions for Polymers; Final Rule” found at 40 CFR Part 723, (60) FR 16316–16336, Published Wednesday, March 29, 1995. Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency, Washington, DC.
- Wambaugh, J.F., Wetmore, B.A., Pearce, R., Strobe, C., Goldsmith, R., Sluka, J.P., Sedykh, A., Tropsha, A., Bosgra, S., Shah, I., Judson, R., Thomas, R.S., Setzer, R.W., 2015. Toxicokinetic triage for environmental chemicals. *Toxicol. Sci.: Off. J. Soc. Toxicol.* 147, 55–67. <https://doi.org/10.1093/toxsci/kfv118>.
- Wetmore, B.A., Wambaugh, J.F., Allen, B., Ferguson, S.S., Sochaski, M.A., Setzer, R.W., Houck, K.A., Strobe, C.L., Cantwell, K., Judson, R.S., LeCluyse, E., Clewell, H.J., Thomas, R.S., Andersen, M.E., 2015. Incorporating high-throughput exposure predictions with dosimetry-adjusted in vitro bioactivity to inform chemical toxicity testing. *Toxicol. Sci.: Off. J. Soc. Toxicol.* 148, 121–136. <https://doi.org/10.1093/toxsci/kfv171>.
- Wetmore, B.A., Wambaugh, J.F., Ferguson, S.S., Sochaski, M.A., Rotroff, D.M., Freeman, K., Clewell, H.J., Dix, D.J., Andersen, M.E., Houck, K.A., Allen, B., Judson, R.S., Singh, R., Kavlock, R.J., Richard, A.M., Thomas, R.S., 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol. Sci.: Off. J. Soc. Toxicol.* 125, 157–174. <https://doi.org/10.1093/toxsci/kfr254>.
- Willmann, S., Lippert, J., Sevestre, M., Solodenko, J., Fois, F., Schmitt, W., 2003. PK-Sim®: a physiologically based pharmacokinetic ‘whole-body’ model. *Biosilico* 1, 121–124. [https://doi.org/10.1016/S1478-5382\(03\)02342-4](https://doi.org/10.1016/S1478-5382(03)02342-4).
- Worth, A.P., Berggren, E., 2024. A Twin Transition in Regulatory Toxicology-moving towards Chemicals 2.0 and phasing out animal testing. *Toxicological sciences: an official journal of the Society of Toxicology*. <https://doi.org/10.1093/toxsci/kfae130>.
- Worth, A.P., Berggren, E., Prieto, P., 2024. Chemicals 2.0 and why we need to bypass the gold standard in regulatory toxicology. *Altern. Lab. Anim. ATLA*, 2611929241296328. <https://doi.org/10.1177/02611929241296328>.