



LJMU Research Online

Thompson, C, Firman, JW, Goldsmith, M, Grulke, C, Tan, Y-M, Paini, A, Penson, P, Sayre, R, Webb, S and Madden, JC

A systematic review of published physiologically-based kinetic models and an assessment of their chemical space coverage

<http://researchonline.ljmu.ac.uk/id/eprint/15634/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Thompson, C, Firman, JW, Goldsmith, M, Grulke, C, Tan, Y-M, Paini, A, Penson, P, Sayre, R, Webb, S and Madden, JC A systematic review of published physiologically-based kinetic models and an assessment of their chemical space coverage. Alternatives to Laboratory Animals. ISSN 0261-

LJMU has developed [LJMU Research Online](http://researchonline.ljmu.ac.uk) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

A systematic review of published physiologically-based kinetic models and an assessment of their chemical space coverage

Courtney V Thompson¹, James W Firman¹, Michael R Goldsmith², Christopher M Grulke², Yu-Mei Tan³, Alicia Paini⁴, Peter E Penson¹, Risa R Sayre², Steven Webb⁵ and Judith C Madden^{1*}

¹School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

²U.S. Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure, Chemical Characterization and Exposure Division, 109 TW Alexander Dr, Research Triangle Park, NC 27709, USA

³U.S. Environmental Protection Agency, Office of Pesticide Programs, Health Effects Division, 109 TW Alexander Dr, Research Triangle Park, NC 27709, USA

⁴European Commission Joint Research Centre (JRC), Ispra, VA, Italy

⁵Syngenta, Product Safety, Early Stage Research, Jealott's Hill International Research Centre, Bracknell, RG42 6EY, England

**Corresponding author: j.madden@ljmu.ac.uk*

Abstract

Across multiple sectors, including food, cosmetics and pharmaceutical industries, there is a need to predict the potential effects of xenobiotics. These effects are determined by the intrinsic ability of the substance, or its derivatives, to interact with the biological system, and its concentration-time profile at the target site. Physiologically-based kinetic (PBK) models can predict organ-level concentration-time profiles, however, the models are time and resource intensive to generate *de novo*. Read-across is an approach used to reduce or replace animal testing, wherein information from a data-rich chemical is used to make predictions for a data-poor chemical. The recent increase in published PBK models presents the opportunity to use a read-across approach for PBK modelling, i.e. to use PBK model information from one chemical to inform the development or evaluation of a PBK model for a similar chemical. Essential to this process, is identifying for which chemicals a PBK model already exists. Herein, the results of a systematic review of existing PBK models, compliant with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) format, are presented. Model information including species, sex, life-stage, route of administration, software platform used and availability of model equations was captured for 7,541 PBK models. Chemical information (identifiers and physico-chemical properties) has also been recorded for 1,150 unique chemicals associated with these models. This PBK model dataset has been made readily accessible, as a Microsoft Excel spreadsheet, providing a valuable resource, for those developing, using or evaluating PBK models in industry, academia and the regulatory sectors.

Keywords: PBK, PBPK, PBTk, systematic review, pharmacokinetic modelling; read-across

Introduction

Humans, like other animals, are exposed daily to a multitude of chemicals of anthropogenic origin, including pharmaceuticals, food additives, pesticides, consumer goods and cosmetic ingredients. The safety assessment of chemicals is a legal requirement and essential to ensure safe use for workers and consumers, as well as protection of domestic/farm animals and environmental species. However, for the majority of chemicals there is a lack of available data for safety assessment, hence predictive models are essential. Predicting toxicity requires knowledge of both the intrinsic activity of the chemical (or its derivatives) and the extent to which the organism is exposed. Whilst external exposure, or dose, has traditionally been used in assessments, it is recognised that the dose at the target site, i.e. organ-level exposure, is the more relevant measure, being causally linked to observed toxicity.¹ Similar reasoning applies in product design where the internal exposure level is linked with a desirable (e.g. therapeutic) effect. In terms of legislation, kinetic data is a specific requirement for plant protection and biocidal product safety assessment and, whilst not formally required, incorporation of such data is widely recommended in other regulations, such as for classification, labelling and packaging (CLP) and the Registration, Evaluation Authorisation and restriction of Chemicals (REACH).² Guidance documents from the European Chemicals Agency³ and the Scientific Committee on Consumer Safety⁴ recommend making use of all available data (including kinetic data) to support decision-making. Whilst general information regarding absorption, distribution, metabolism or excretion (ADME) may be useful, more accurate prediction requires organ-level concentration-time profiles. Physiologically-based kinetic (PBK) models (synonymous with

physiologically-based pharmacokinetic, toxicokinetic or biokinetic (PBPK, PBTK or PBBK) models) are employed in numerous industries to provide such predictions.

In a PBK model the body is represented as a series of compartments (for example individual organs) connected by blood flow. The models use knowledge of physiology and anatomy (such as organ volumes and cardiac output), in combination with chemical-specific information (such as solubility and partitioning behaviour) to predict the concentration time-profile of the chemical in tissues, cellular compartments or sub-compartments.

Differential equations are used to describe the rate of change of concentration of the chemical in each compartment, as summarised in Figure 1. Detailed information on how to construct and validate PBK models, their applications in different sectors and tools available to support PBK modelling have been well-reported previously.^{1, 5-9} Of particular note, is the recent Organisation for Economic Cooperation and Development (OECD) guidance on the characterisation, validation and reporting of models for regulatory purposes.¹⁰ This document builds on the principles described in the WHO report of 2010¹ but focuses on the use of alternative approaches (*in silico* and *in vitro*) for parameterising PBK models. The potential for applying new approach methodologies (NAMs) and next generation (NG) methods to support the development and use of PBK models in safety assessment, was also promulgated by Paini and colleagues.¹¹ PBK models can assimilate new information as it becomes available to increase predictive capacity; these models provide an advantage over traditional one or two compartment kinetic models.¹²

For environmental chemicals, the numerous applications of these models include: determining the dose at target tissue following external exposure; route-to-route extrapolation; dose extrapolation; inter- and intra-species extrapolation (accounting for species, population or genetic variability through adaptation of physiological and anatomical parameters); *in vitro*-to-*in vivo* extrapolation (IVIVE); ascertaining safe levels based on tissue dosimetry; estimating chemical exposure from biomonitoring or epidemiological data (using reverse dosimetry) and assessing potential for bioaccumulation. These applications complement the traditional role of PBPK modelling of drugs where they can be utilised to predict first dose in man, potential for drug-drug interactions or the influence of health status (e.g. hepatic impairment) on kinetics.¹³

ECHA reports that read-across is the most commonly used alternative method to reduce or replace animal testing in safety assessment.¹⁴ In this approach, information from a data-rich (source) chemical is used to predict information for a data-poor (target) substance that is considered similar.³ Kinetic information plays a key role in supporting read-across predictions^{3,15} and recent efforts have aimed to increase the accessibility of such data. Sayre and co-workers¹⁶ published a database of time-series concentration data, extracted from an extensive search of the literature, and Pawar and co-workers identified 38 databases containing a range of ADME-relevant data as part of their overall review of resources to support read-across and *in silico* model development.¹⁷ PBK models provide an additional opportunity to derive data to support read-across. Data may be acquired either from a PBK model for the chemical under investigation (the target chemical) or from a PBK model for an existing chemical considered similar to the target (a source chemical). This latter approach,

wherein an existing PBK model for a source chemical is used as a template for a target chemical, is contingent upon the identification of existing, suitable PBK models.

Over the past 30 years the number of published PBK models and their applications has increased significantly.¹⁸ In 2016 Lu and co-workers published a PBK Knowledgebase, comprising 307 chemicals for which PBK models were available from papers published between 1977 and 2014.¹⁹ In their report the authors described two case studies, wherein PBK models from the Knowledgebase were used to inform the development of PBK models for “similar” chemicals. In their study chemical analogues were identified based on similarity of physico-chemical properties, although it is recognised that there is no consensus as to the best method to determine similarity.²⁰ Ellison and Wu successfully demonstrated an analogous approach wherein a PBK model for a target chemical was evaluated using information from source chemicals identified as structural or functional analogues.²¹ In order to assist researchers in identifying existing PBK models, a spreadsheet of those collated from the literature by the US Environmental Protection Agency (EPA), was made available via Figshare.²² This resource included information on species, gender, life-stage, route of administration, compartments and PubMed ID for the source of the models.

An enriched version of this PBK Knowledgebase was recently used as a proof-of-principle to demonstrate that information from an existing PBK model could be used, in a read-across approach, to inform safety assessment.²³ In the analysis, methyleugenol was considered as a target chemical with estragole and safrole being identified as suitable source chemicals (with respect to structural similarity). This approach was also successful, exemplifying how information from an existing PBK model could assist the development of a model for a similar chemical. Making best use of existing data and in particular the application of the

read-across approach are recognised as important tools in reducing animal testing.²⁴ In order to facilitate the application of this approach, it is essential to identify chemicals for which PBK models are available. As “similarity” is often considered in relation to structure or physico-chemical properties it is also important to ascertain the nature of the chemicals for which models are available, comparing their characteristics to existing chemical datasets. Having information regarding the chemicals and the models in a readily accessible and updateable resource would be a significant asset for researchers, industry and regulators, with the potential to reduce the number of animals used in drug development and chemical safety assessment.

Several key features characterise an individual PBK model, these are represented in Figure 1 and include species, sex, life-stage, route of administration and the compartments required to accurately describe the time-course of the chemical. In some models, key organs (such as liver, lung etc) are incorporated individually as compartments, in others these are further divided into constituent sub-compartments (for example considering histopathological regions or explicitly including lymph or interstitial / vascular space) giving higher-level, more complex models. In other scenarios, organs are grouped together (referred to as “lumping”) to create simpler models, for example all poorly perfused organs are considered as one compartment and all highly-perfused organs are considered as another. In addition to the physiological and anatomical information required, chemical-specific data is also a pre-requisite. A substance may be identified using common names or chemical identifiers such as the Chemical Abstracts Registry (CAS) number, a Simplified Molecular Input Line Entry System (SMILES) string or the International Chemical Identifier Key (InChiKey). Ideally multiple identifiers should be incorporated in the dataset to avoid ambiguity. Model

development can be performed using a range of software, the equations employed may be specified within the publication itself (or as part of the supplementary information accompanying the article). Within this systematic review, key model characteristics, such as species, sex, life-stage, route of administration, compartments, availability of model equations and chemical identifiers were captured within the PBK model dataset, as summarised in Figure 1.

The second part of the analysis relates to the assessment of the chemical space coverage of the PBK model dataset. There is no simple process by which a chemical can be designated as being a particular “type” – for example cosmetic ingredients may also be food additives; botanicals may have pharmaceutical properties etc. Consequently, in order to assess the nature of the chemicals in the PBK model dataset, key physico-chemical properties were generated and compared to those of chemicals appearing in other datasets. Datasets studied were botanicals, pesticides, pharmaceuticals, food, cosmetic ingredients and REACH chemicals. The number of chemicals in the PBK model dataset that also appeared in each of the other datasets was ascertained.

In summary, the aim of this systematic review was the curation of a data resource for existing PBK models. Relevant information for the models (species, sex, life-stage, substance identity, software used etc) was captured in a flexible spreadsheet format. The chemical space occupied by PBK models (in terms of physico-chemical properties) was compared to that of other chemical types, using six existing datasets. This resource has been created to assist the development and evaluation of PBK models using existing data, thereby reducing the need to generate new data from animal studies.

[FIGURE 1 NEAR HERE]

Methods

Systematic Review

This systematic review was prospectively registered on PROSPERO, the National Institute for Health Research's international prospective registration system with the review question stipulated as: "For which substances are physiologically-based kinetic (PBK) models available and which species, genders, life-stages and routes of administration have been investigated for these substances? This will include determining the chemical space coverage of the models and the availability of the associated model equations within the literature."²⁵ The review complies with the PRISMA reporting standards - the PRISMA checklist is available as Supplementary Material (iii).

Briefly, following a scoping study of potentially useful databases and search terms, Scopus (<https://www.scopus.com/>), PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) and Web of Science (<https://www.webofknowledge.com>) were selected as the most appropriate databases for identifying published papers on PBK models. The search of these databases was completed in October 2020. The search terms ("pbpk" OR "pbk" OR "pbbk" OR "pbtok" OR "pbpd" OR "pbpm" OR "physiologically based") AND ("pharmacokinetic" OR "toxicokinetic" OR "biokinetic" OR "pharmacodynamics" OR "biopharmaceutical") were used to search abstracts, titles and keywords of papers within each database, across all years available. The systematic review management tool Covidence was used for processing

papers for the review (<https://www.covidence.org/>, accessed May 2021). A total of 14,803 papers were initially identified, however following automated removal of duplicates in Covidence, 6,771 remained. All abstracts were screened independently by two researchers with all conflicts being resolved by discussion. The inclusion criteria encompassed PBK models for all routes of administration for chemical, biological and carrier systems, including cases where normal physiology was altered or interactions between administered substances were investigated. Models that could not be associated with a specific substance (such as generic models applicable to large groups of compounds) were excluded. Where an abstract was associated with a paper that had subsequently been retracted it was ensured that this model was excluded from the dataset. Although standard practice in other systematic reviews, assessment of the quality of the reported models and risk of bias in reporting was considered unnecessary for this review. Our intention here was to document all available models, enabling interested researchers to rapidly identify potentially useful models to assist with future model development. The assessment of PBK model quality (aside from fundamental considerations relating to good modelling practice) needs to be considered in terms of fitness for a given purpose,¹⁰ it is therefore context dependent and remains the decision of the model user. Following abstract screening, 3,120 abstracts were retained for full text screening. PBK model data were extracted from 1,649 of these papers, resulting in 7,541 individual models being captured. Note that if oral and intravenous dosing were used for both male and female subjects for the same chemical this would be extracted as four individual models, hence there are many more models than individual chemicals. Reasons for exclusion of papers during full text screening included: PBK model not being reported in the article; the article referenced a previously published model with no

adaptations (information on the PBK model was extracted from the original publication); full article not being available in English or not being reasonably accessible.

Extraction of data from available PBK models

Data were manually extracted from these 1,649 papers by one reviewer, with information being acquired from text, tables, figures and supplementary information. The data were entered into a Microsoft Excel spreadsheet that captured details of the chemical: parent and metabolites specified where appropriate, species (with sub-category where relevant), sex, life-stage, route of administration, literature reference for the model (with DOI), compartments considered in the model, the software employed and the availability of PBK model equations within the article. Controlled vocabulary was used where possible to ensure consistency of data extraction and to enable the resulting spreadsheet to be readily filtered and searched for specific types of models, i.e. controlled vocabulary was used for species, sex, life-stage, route of administration, availability of equations and software used. The vocabulary was empirically derived, to enable the most efficient searching, the full rationale is given in the Supplementary Material (i). For example, life stages can be reported in multiple ways – no. of weeks, months, years of age, young adult, adult, neonate, young child, juvenile etc. Therefore, for consistency this information was allocated to the more generic categories of pre-birth or pre-hatch; from birth or hatch up to adult; adult; pregnant; old age (if specified) as well as a generic category for health compromised (excluding old age) individuals.

Chemical identifier information was obtained by manually inputting the chemical name (as given in the publication) into PubChem (<https://pubchem.ncbi.nlm.nih.gov/>; last accessed May 2021) and extracting the molecular weight, canonical SMILES, isomeric SMILES, InChiKey and CAS number. The CAS registry number from PubChem was used as input for the COSMOS database, version 2 (<https://cosmosdb.eu>, accessed April 2021). Where available, the CAS registry number and chemical name, as recorded in COSMOS, were extracted to confirm the identity of the chemical; the COSMOS ID was also extracted. This information was captured by one reviewer. An assessment of the reliability of screening and data extraction was undertaken and is reported below.

Assessment of the chemical space coverage of the PBK model dataset in relation to other chemical datasets

In order to identify the nature of the chemicals in the PBK model dataset and to compare the chemical space coverage (in terms of key physico-chemical properties) six existing datasets were investigated. These have been compiled in-house at Liverpool John Moores University and comprise: botanicals; pesticides; cosmetic ingredients (obtained from COSMOS db version 2 (<https://cosmosdb.eu>; accessed April 2021); food additives (obtained from <http://foodb.ca>; accessed April 2021); pharmaceuticals (obtained from www.drugbank.ca; accessed April 2021) and REACH chemicals as summarised in Table 1.

Canonical SMILES for all chemicals in these datasets were generated using OpenBabel (v.3.0.0; http://openbabel.org/wiki/Main_Page; accessed April 2021). From the PBK model dataset 1,150 unique SMILES were identified with 1,187 unique InChiKeys (note that chemical isomers may have the same SMILES string but different InChiKeys). In order to

determine how many chemicals with PBK models were present in each of the other six datasets the InChiKeys were compared.

The SMILES strings for all datasets were inputted into the RDKit (v. 2020.03.6; www.rdkit.org) Descriptor Node, accessed through KNIME software (v. 4.3.1; www.knime.com), in order to obtain the physico-chemical properties for all chemicals. Properties included molecular weight, number of hydrogen bond donors / acceptors, predicted logarithm of the octanol:water partition coefficient (SlogP) and the topological polar surface area (TPSA); the number of Lipinski rule violations were calculated from this information. Whilst it is possible to generate thousands of physico-chemical properties, here only a few readily calculable properties were selected, representing those most often used to broadly characterise chemicals in terms of size, polarity and partitioning behaviour. These simple properties were also used to determine Lipinski rule violations, frequently used to indicate potential oral absorption – a common route of administration for these models. The minimum, maximum, mean and median values and interquartile ranges of these properties were calculated using Minitab version 19.2 for all datasets. Histograms were also generated using Microsoft Excel to enable a visual comparison of the property ranges between the different datasets. The results of the statistical analysis are available as Supplementary Material (ii).

Results

Systematic Review

Figure 2 summarises the number of papers considered at each stage of the review process and the final number of models extracted in the PBK model spreadsheet. Of the 6,771 of papers initially identified, 3,120 remained after abstract screening and data were extracted from 1,649 of these.

[FIGURE 2 NEAR HERE]

Extraction of data from available PBK models

A total of 7,541 individual models were identified and extracted. Note that for an individual substance multiple models may be available in one paper. For example, if male and female subjects were used and doses given intravenously and by oral dosing this would be classed as four models, this enables more flexible searching; searches can be conducted, not just by chemical names, but also by species or route of administration etc. A total of 1,889 chemical names are present in the PBK model dataset, some represent biological entities (such as monoclonal antibodies) or are not associated with a specific structure, resulting in 1,187 unique InChiKeys, i.e. unique chemicals with identifiable structures. Summary information concerning the PBK models extracted is shown in Figures 3 and 4. Figure 3 shows the

breakdown of models by species investigated and Figure 4 shows the breakdown by route of administration.

[FIGURE 3 NEAR HERE]

[FIGURE 4 NEAR HERE]

Assessment of the chemical space coverage of the PBK model dataset in relation to other chemical datasets

The results for the comparison of InChiKeys for chemicals in the PBK model dataset to those for the six comparative datasets are shown in Table 1.

Table 1. The number (and percentage) of chemicals with existing PBK models that are present in the six comparative datasets investigated.

Dataset	Number of chemicals in dataset (with unique, identifiable structure)	Number of chemicals in dataset also present in PBK model dataset	Percentage of chemicals in dataset also present in PBK model dataset (%)
PBK model	1,186	N/A	N/A
Botanical	947	24	2.67
Pesticide	945	40	4.23
COSMOS	5,655	123	2.41
Food	2,674	88	3.37
DrugBank	6,716	382	5.80
REACH	75,663	620	0.85

Table 2 shows the results of the statistical analysis of the key physico-chemical properties (molecular weight, number of hydrogen bond donors and acceptors, logarithm of the octanol:water partition coefficient, topological surface area and number of Lipinski rule violations for the chemicals in the PBK model dataset.

Table 2. Statistical analysis of the physico-chemical properties of the chemicals in the PBK model dataset.

Variable ^a	Mean	Minimum	Q1	Median	Q3	Maximum	Range	IQR
MW	325.61	6.94	163.10	292.28	410.67	6496.20 ^b	6489.26	247.57
No. HBD	1.79	0.00	0.00	1.00	2.00	100.00	100.00	2.00
No. HBA	4.17	0.00	1.00	3.00	6.00	120.00	120.00	5.00
SlogP	2.25	-45.03 ^b	0.80	2.08	3.73	11.10	56.13	2.93
TPSA	73.07	0.00	23.47	56.92	93.00	2536.36	2536.36	69.53
nViolations	0.37	0.00	0.00	0.00	0.00	4.00	4.00	0.00

**MW = molecular weight; No. HBD = number of hydrogen bond donors; No. HBA = number of hydrogen bond acceptors; SlogP = predicted logarithm of the octanol:water partition coefficient; TPSA = topological polar surface area; nViolations = number of violations of the Lipinski Rule of Five (Lipinski et al 1997).*

^bThe extreme values here are for vistarem a magnetic resonance imaging contrast agent with large hydrophilic chains

Supplementary Material (ii) shows the results of a similar statistical analysis of the physico-chemical properties for the six datasets to which the PBK model dataset was compared.

Figure 5 (A-F) shows the comparison of these key physico-chemical properties across the seven datasets.

[FIGURE 5 NEAR HERE]

Assessment of screening and data extraction reliability

After screening all 6,771 abstracts in duplicate, 3,120 were taken forward to full text screening; of these, 1,362 papers were rejected at this stage. In addition, 109 papers could not be readily obtained or were not in English, therefore data were extracted from 1,649 papers, resulting in 7,541 models. The resultant spreadsheet comprises over 150,000 individual entries as for each model the species, sex, life-stage, route of administration, availability of equations, compartments, references and chemical identifiers were captured. It is expected that errors will arise when assessing the suitability of papers for inclusion and performing extensive manual processing, hence a quality assessment exercise was

undertaken. Three aspects of the screening and data extraction processes were reviewed by a second investigator:

(i) 5 % of the papers that had been excluded at the full text screening phase were reviewed.

(ii) 5 % of PBK model data extracted from the papers (chemical information, species data (primary and secondary categories), sex, life-stage, route of administration, reference (DOI and PubMed ID if available), compartments investigated, availability of equations and simulation software were checked.

(iii) 5 % of the chemical identifier information from PubChem and COSMOS (chemical name, CAS registry number, molecular weight, canonical SMILES, isomeric SMILES, InChiKey, COSMOS ID) was obtained again from these sources and compared to the values in the spreadsheet.

The greatest source of “error” was determined to be the exclusion of papers that were considered as potentially relevant by a second investigator i.e. 6% of excluded papers. In terms of the systematic review this is not considered a highly significant problem. PBK models are continually being published, hence there can never be a finalised set of models. It is the intention to make this resource available in its current form, as a tool to assist researchers in finding relevant PBK models, and to update the resource in future capturing models previously not identified or erroneously excluded.

For PBK model data, manually extracted from the papers, an error was detected in the information captured for 3.5 % of substances. This does not equate to 3.5 % of the total

information being incorrect as this may indicate an error in only one (or possibly more) of the 13 columns that relate to the PBK model information.

An error was detected in the data for 2.4% of the chemicals in relation to the identifier information. As above, this does not equate to 2.4% of the total information being incorrect, but that for 2.4% of chemicals an error was detected in one (or more) of the 7 columns associated with chemical identifier information.

The authors welcome any feedback from users regarding errors, omissions of existing models or updates for new models (note that models require a minimum level of information and novelty to be included) please e-mail the corresponding author.

Discussion

In this systematic review, information concerning over 7,500 PBK models were extracted from 1,649 papers. The models encompassed 18 species (including rat, human, mouse, cow, guinea-pig) at various life-stages (e.g. juvenile, adult, pregnant, health-compromised) across 21 administration routes (e.g. oral, inhalation, in utero). The information has been distilled into a Microsoft Excel spreadsheet that was constructed using controlled vocabulary to enable users to search using different criteria (e.g. to select models by species or routes of administration etc). It is anticipated that researchers or regulatory scientists can use this information to assist building or evaluation of new models, or as a resource from which to extract relevant pharmacokinetic or toxicokinetic data.

Although this is the largest collation of PBK models that the authors are aware of, it is not a complete list. As identified in the quality assessment exercise, some of the historic models were omitted, also it would not be possible to generate a finite list of all models as this is such a dynamic area of research. The publication of new models has shown a rapidly increasing trajectory in recent years,²⁶ the current dataset serves as a basis for continuing curation of existing models which will provide an increasingly rich source of information for modellers in future.

Chemicals can be used for a variety of purposes and it is often not feasible to allocate a chemical to unique group (e.g. there is a significant cross-over between chemicals used as food additives and as cosmetic ingredients, hence the same chemicals may appear in more than one of the different datasets). It is therefore difficult to determine for which chemical “types” there are most PBK models, however there are some trends discernible from the analysis undertaken. Unsurprisingly, given that PBK modelling evolved in drug development, the greatest proportion of models correspond to chemicals in the DrugBank dataset.

Pesticides are generally well studied and data-rich, therefore the second most common type of chemical with PBK models are the pesticides. For food additives and cosmetic ingredients, where there are often chemicals in common, similar proportions of chemicals have PBK models. Due to the size and generality of the REACH dataset it would be anticipated that relatively few chemicals would have existing PBK models. The results confirm the paucity of PBK models available in relation to different areas of chemical space and underline the importance of using existing PBK models to help fill data gaps.

Figure 5.A shows the distribution of molecular weight across the seven datasets. As expected, the majority of chemicals fall within the range 100-600 Da but there are notable

differences between the datasets. There are a relatively high number of chemicals in the PBK dataset with low molecular weight, these will include the volatile chemicals for which respiratory uptake has been extensively studied. Food additives and cosmetic ingredients (which have chemicals in common) show a relatively high proportion of low molecular weight chemicals. Chemicals that are designed to be biologically active, such as drugs and pesticides, tend to be developed in accordance with guidelines relating to preferred physico-chemical properties. For example, the Lipinski Rule of Five stipulates that poor oral absorption of drugs is associated with those chemicals having molecular weight > 500 Da; $\log P > 5$; >10 hydrogen bond acceptors or; > 5 hydrogen bond donors.²⁷ Other research has also suggested that $TPSA > 140 \text{ \AA}^2$ is also unfavourable for oral absorption, consequently certain chemical types are designed to fall within narrower property ranges and such trends are evident in the property ranges here. A correlation between molecular weight and $\log P$ is often observed amongst groups of chemicals (although there are many exceptions to this), here a similar pattern to the range of values is generally observed for $\log P$ and molecular weight. Pesticide and botanical datasets have a greater percentage of chemicals in the $\log P$ ranges 3-4 (43% and 47% respectively), whereas the PBK dataset only has 28% of chemicals in this range. Pesticide and botanical datasets comprise fewer molecules capable of carrying a charge (associated with increased hydrophilicity) hence on average have higher $\log P$ values. This is significant as partitioning behaviour (often estimated using $\log P$) is a key element in building PBK models. Whilst the extreme values for $\log P$, calculated by the software used here, may be unrealistic (and therefore unsuitable for model building), when used for comparison they are still useful for demonstrating the trends in the data. The range in values for all of the physico-chemical properties of pesticides, is narrower than for the other chemical types, indicating the more restrictive chemical properties required for these

chemicals. Botanicals generally show a wider range of values for each of the physico-chemical properties (in particular more chemicals show properties at the upper extremes of the ranges). A significant number of compounds within this set are large and complex. Whilst in other datasets, molecules tend to be designed for a specific purpose (and those outwith given property ranges are filtered out) the same exclusions would not be applicable to this dataset.

Conclusion

Understanding the kinetic behaviour of a chemical within the body, particularly its concentration-time profile at a target site, is essential to accurately determine its potential effect. For the majority of chemicals there is a lack of data concerning toxicity and kinetics, however generating such information *de novo* would require excessive use of animals and is legally, ethically and financially constrained. Hence there is a need to leverage existing knowledge in order to obtain as much information as possible to assist decision making. Read-across is the most common method used by which information from data-rich chemicals is used to predict information for data-poor chemicals. Herein is presented a comprehensive collation of existing PBK models that can be searched using multiple criteria. The physico-chemical space of PBK models has been mapped against that of other chemical types ie. food additives, cosmetic ingredients, drugs, REACH chemicals, botanicals and pesticides. Organising the current state of knowledge of existing PBK models provides a valuable resource for those working in the area to identify models for chemicals of interest, or analogues, that can be used to assist the development or evaluation of new PBK models.

The concept of using existing PBK model information in a read-across approach to develop a new PBK model for an analogous chemical has been demonstrated successfully in recent publications.^{18,22} Such an approach is recommended in the recent OECD guidance on PBK modelling which focuses on the use of alternative methods in PBK model development.¹⁰ The PBK model dataset described herein enables researchers to readily gain insight into available PBK models across multiple species, life-stages and routes of administration, such that the structure and parameterisation of PBK models for different chemicals is more accessible. This ensures maximum use of existing knowledge on PBK modelling and reduces the time and cost associated with developing new PBK models.

Continuing effort is required to curate existing PBK models. The ability to extract relevant data from models and reproduce those published in literature would be facilitated by researchers embracing the use of systematic methods to record PBK models. Consistent formats for publishing PBK models have been proposed previously,^{28,29} with the template proposed in the recent OECD guidance¹⁰ drawing on previous recommendations, most significantly the reporting format proposed by Tan and co-workers³⁰. Use of consistent reporting formats is strongly encouraged as this assists other researchers in re-using or re-purposing existing models.

Future effort is required to curate available models, identify appropriate similarity metrics to assist in the identification of PBK models for analogues to chemicals of interest and, ideally, to make such resources widely available, for example by incorporation into free webtools. Such endeavours will help to leverage the wealth of information available in existing PBK models.

Supporting Information

Supporting information to accompany this manuscript includes:

Supplementary material (i) a Microsoft Excel file incorporating the spreadsheet used to capture model information for over 7,500 models, an explanation of the controlled vocabulary employed, and the physico-chemical properties of the chemicals for which a PBK model has been identified.

Supplementary material (ii) full results of the statistical analysis of the chemical space of the datasets investigated.

Supplementary material (iii) PRISMA compliance checklist for systematic reviews

The Endnote library of papers from which the model data were extracted is available on request; references / DOIs are available in the Microsoft Excel sheet in Supplementary Material (i).

Declaration of conflicting interests

The authors declare no conflict of interest with respect to the research, authorship or publication of this article.

Ethical approval

Ethics approval was not required for this article.

Informed Consent

Informed consent was not required for this article

Funding

Judith Madden and Courtney Thompson gratefully acknowledge the funding and scientific advice of the European Partnership for Alternative Approaches to animal testing (EPAA) with respect to the funding of a studentship for Courtney Thompson.

Disclaimer

The views expressed in this article are those of the author(s) and do not necessarily represent the views or the policies of the U.S. Environmental Protection Agency.

References

[1] WHO, (2010) Characterisation and application of physiologically based pharmacokinetic models in risk assessment. Available at: <https://apps.who.int/iris/handle/10665/44495>.

(Accessed May 2021).

[2] Pistollato F, Madia F, Corvi R et al. Current EU Regulatory requirements for the assessment of chemicals and cosmetic products: challenges and opportunities for

introducing new approach methodologies. *Arch Toxicol* 2021;

<https://doi.org/10.1007/s00204-021-03034-y>

[3] European Chemicals Agency (ECHA, 2017) Read-across assessment framework (RAAF)

DOI: 10.2823/619212

[4] Scientific Committee on Consumer Safety (SCCS) Notes of guidance for the testing of cosmetic ingredients and their safety evaluation, 2018. 10th revision (SCCS/1602/18).

[5] US EPA (2006). Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment (final report). U.S. Environmental Protection Agency, Washington, D.C., EPA/600/R-05/043F, 2006.

[6] Kuepfer L, Niederal C, Wendl T. et al. Applied concepts in PBPK modeling: how to build a PBPK/PD model. *CPT Pharmacometrics Syst Pharmacol*, 2016; 5: 516-531.

[7] Upton RN, Foster DJ and Abuhelwa AY. An introduction to physiologically-based pharmacokinetic models. *Paediatr Anaesth*, 2016; 26: 1036-1046.

[8] Laroche C, Aggarwal M, Bender H et al. Finding synergies for 3Rs - Toxicokinetics and read-across: report from an EPAA partners' Forum. *Regul Toxicol Pharmacol*, 2018; 99; 5-21.

- [9] Madden JC, Pawar G, Cronin MTD et al. In silico resources to assist in the development and evaluation of physiologically-based kinetic models. *Comput Toxicol* 2019; 11; 33-49.
- [10] OECD (2021), Guidance document on the characterisation, validation and reporting of Physiologically Based Kinetic (PBK) models for regulatory purposes, OECD Series on Testing and Assessment, No. 331, Environment, Health and Safety, Environment Directorate, OECD, Paris
- [11] Paini A, Leonard JA, Joossens E et al. Next generation physiologically based kinetic (NG-PBK) models in support of regulatory decision making. *Comput Toxicol* 2019; 9: 61-72.
- [12] Rowland M, Lesko LJ and Rostami-Hodjegan A. Physiologically Based Pharmacokinetics Is Impacting Drug Development and Regulatory Decision Making. *CPT Pharmacometrics Syst Pharmacol* 2015; 4: 313-315.
- [13] Perry C, Davis G, Conner TM et al. Utilization of Physiologically Based Pharmacokinetic Modeling in Clinical Pharmacology and Therapeutics: an Overview, *Curr Pharmacol Rep* 2020; 6; 71–84.
- [14] ECHA (2020). The use of alternatives to testing on animals for the REACH Regulation. Fourth report under Article 117(3) of the REACH Regulation. Helsinki: European Chemicals Agency, 2020, 85 pp

[15] Schultz TW, Amcoff P, Berggren E et al. A strategy for structuring and reporting a read-across prediction of toxicity. *Regul Toxicol Pharmacol* 2015; 72: 586–601

[16] Sayre RR, Wambaugh JF and Grulke C. Database of pharmacokinetic time-series data and parameters for 144 environmental chemicals. *Scientific Data* 2020; 7:122

<https://doi.org/10.1038/s41597-020-0455-1> (Accessed May 2021)

[17] Pawar G, Madden JC, Ebbrell, D et al. In Silico Toxicology Data Resources to Support Read-Across and (Q)SAR. *Front Pharmacol* 2019 doi: 10.3389/fphar.2019.00561

[18] Tan Y-M, Worley RR, Leonard JA and Fisher JW. Challenges Associated With Applying Physiologically Based Pharmacokinetic Modeling for Public Health Decision-Making. *Toxicol Sci*, 2018; 162; 341-348.

[19] Lu J, Goldsmith MR, Grulke C, et al. Developing a physiologically-based pharmacokinetic model knowledgebase in support of provisional model construction. *PLoS Comput Biol*, 2016; 12, e1004495.

[20] Mellor CL, Marchese Robinson RL, Benigni R, et al. Molecular fingerprint-derived similarity measures for toxicological read-across: recommendations for optimal use. *Regul Toxicol Pharmacol* 2019; 101: 121–134.

[21] Ellison CA and Wu S. Application of structural and functional pharmacokinetic analogs for physiologically based pharmacokinetic model development and evaluation. *Regul Toxicol Pharmacol* 2020; 114: 104667.

[22] Sayre RR, Grulke C, Tan, Y-M et al Existing physiologically-based kinetic (PBK) models, identified via PubMed searches (with associated EndNote library). Available at: https://figshare.com/articles/dataset/Existing_physiologically-based_kinetic_PBK_models_identified_via_PubMed_searches_with_associated_EndNote_library_/10075574, 2019 (Accessed May 2021)

[23] Paini A, Worth A, Kulkarni S et al Assessment of the predictive capacity of a physiologically based kinetic model using a read-across approach. *Comput Toxicol* 2021: Available at: <https://doi.org/10.1016/j.comtox.2021.100159> Accessed May 2021

[24] Madden, JC, Enoch SJ, Paini, AP and Cronin MTD (2020) A review of in silico tools as alternatives to animal testing: principles, resources and applications, *Altern Lab Anim* 2020; 48: 146–172

[25] Thompson C, Madden J and Penson P. Research Protocol Registration: Systematic review to determine the chemical space of existing physiologically-based kinetic (PBK) models PROSPERO 2020 CRD42020171130, 2020; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020171130 (accessed May 2021)

[26] Paini, A. et al. 2017, Investigating the state of physiologically based kinetic modelling practices and challenges associated with gaining regulatory acceptance of model applications, *Regul Toxicol Pharm* 2017; 90: 104–115.

[27] Lipinski CA, Lombardo F, Dominy BW et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliver Rev* 1997; 23: 3-25

[28] EMA (2018) Guideline on the reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation EMA/CHMP/458101/2016, Available at: <https://www.ema.europa.eu/en/reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation>. Accessed April 2020.

[29] US FDA. (2018) Physiologically based pharmacokinetic analyses – Format and content Guidance for Industry. Available at: <https://www.fda.gov/files/drugs/published/Physiologically-Based-Pharmacokinetic-Analyses-%E2%80%94-Format-and-Content-Guidance-for-Industry.pdf>. Accessed May 21

[30] Tan Y-M, Chan M, Chukwudebe A et al. PBPK Model reporting template for chemical risk assessment applications. *Regul Toxicol Pharmacol* 2020; 115; 104691.

