A Method to Quantify Reproducibility in PBPK Model Methods and Results

Risa R. Sayre[1,2], R. Woodrow Setzer[1], John Wambaugh[1], Christopher M. Grulke[1]

1: U.S. Environmental Protection Agency, Office of Research and Development, National Center for Computational Toxicology 2: Oak Ridge Associated Universities

Risa Sayre | ORCiD 0000-0002-6173-8020 | sayre.risa@epa.gov | 919-541-4871

1. Background & Purpose

Physiologically-based pharmacokinetic (PBPK) modeling describes how the amount of an exogenous **chemical** in the body changes over **time**, using values specific to properties of that chemical in a system of equations representing relevant biological processes.

PBPK models can be used to relate *in vitro* toxicity data and *in vivo* toxicities by correlating external doses to compartment concentrations, enabling the extrapolation of biochemical responses in high-throughput target tissue cell assays to predicted points of departure[1].

toxicity testing | **dose-response (PBPK) modeling** | human exposure data | risk contexts

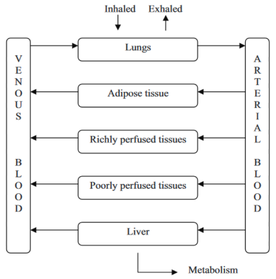


Fig 1. Basic inhalation PBPK model system, showing model compartments and flows (figure from IPSC[5])

Models describe a system by synthesizing knowledge and assumptions into a quantitative expression, so context matters[8]. While building an open PBPK model **database** for risk assessors and modelers from 973 publications, we wanted to effectively capture complete parameterization for all models. We developed a list of criteria to systematically quantify common, fundamental obstacles to **reproduction of methods and results** for this set.

2. Methods

Important concepts in PBPK modeling were gleaned from review documents[2,3,4,5,6] and from an initial review of the model set. From this knowledge, we identified minimal essential, objective, Boolean criteria to represent necessary features for a reproducible PBPK model. Data was sourced from the paper and published supplemental material.

2.1. REPRODUCIBILITY OF METHODS

Documentation:

- ❑ description of the proposed PBPK **system**
- ❑ set of parameter **values** to input into that system
- ❑ time course **results** of application of values to system

Fidelity:

- ❑ complete set of model **equations**
- ❑ dimensional consistency of **units** within those equations
- ❑ preservation of **mass balance** in system flows

2.2 REPRODUCIBILITY OF RESULTS

Support:

- ❑ data presented from *in vivo* experiment
- ❑ measured concentrations at >3 time points
- ❑ data from more than one subject

Code:

- ❑ code included in the paper or supplement
- ❑ code recreates the figures from the paper

$$\frac{dA_t}{dt} = Q_t \times C_{art} - Q_t \times \left(\frac{A_t}{V_t \times P_t}\right) - \left(\frac{V_{max} \times C_t}{K_m + C_t}\right)$$

Fig 2. Equation example (figure from McLanahan et al.[3]). A concentration flows into a compartment at a rate determined by blood flow. The concentration flowing out is reduced by the amount that partitions into the tissue, and the amount that changes into a metabolite, following Michaelis-Menton kinetics.

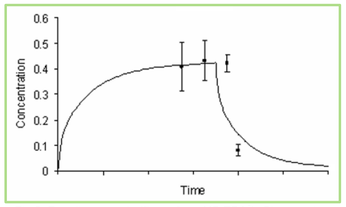
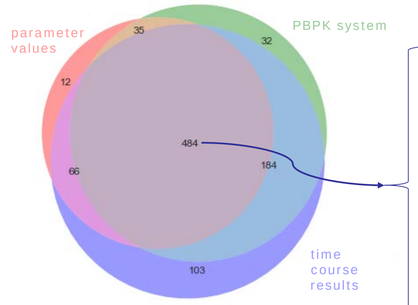


Fig 3. Results example (figure from US EPA[2]). Changes in concentration in a target tissue over time produced by the model is compared to experimental data. More than three data points are required to evaluate fit to a complex line.

3. Results

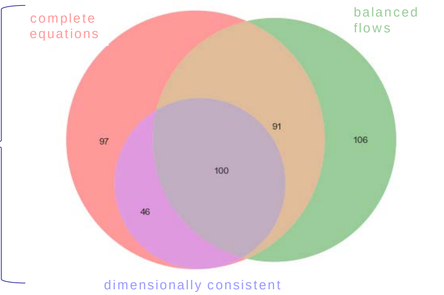
Documentation:

49.7% of the 973 papers had potentially complete documentation. **Parameter values** were the most commonly absent feature (37.5%).



Fidelity:

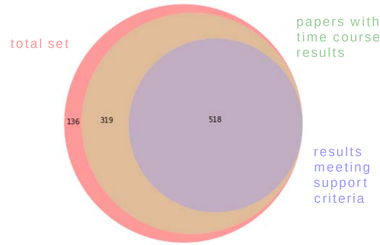
20.7% of papers with complete documentation had an accurate, fully described mathematical basis. The most commonly observed error was inconsistent **units** within the model equations (35.3% were erroneous, and 34.5% of the rest were ambiguous).



3.1: 10.3% of 973 papers had potentially complete documentation and fidelity.

Support:

We examined experimental evidence supporting results for the entire set, regardless of methods reproducibility. The most common observation was a fit to **too few time points** for adequate model validation (15.9%).



Code:

2.6% of the total set of papers provided model code. No published code reproduced result figures from its paper. Inconsistencies between paper and code and errors were noted.

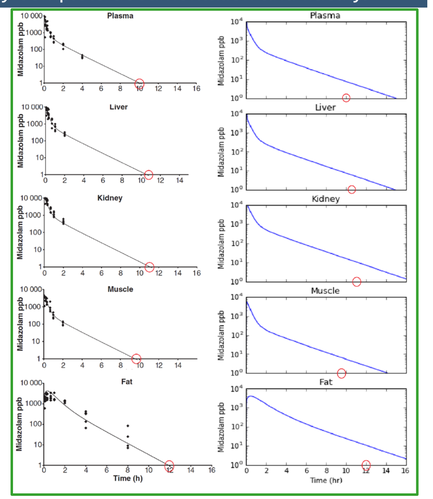


Fig 4. Results figures from paper on left, code implementation on right. Red circles show time to reach 1 ppb in paper (faster than code for all compartments). The significance of the difference is outside our scope.

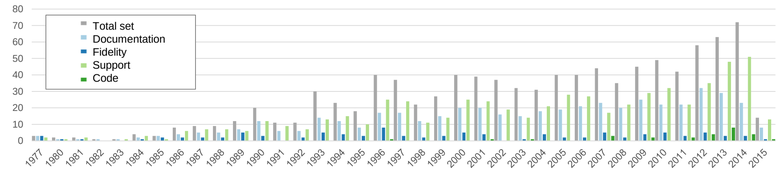
3.2: 53.2% of 973 papers presented potentially adequate data support, and no results were reproducible from code.

4. Conclusions & Discussion

The results demonstrate that the evaluated criteria are common elements missing from PBPK model literature. However, since none of the papers were ultimately reproducible, these criteria are **not a sufficient set** to identify and quantify barriers to reproducibility.

Remaining questions:

- Were the most broad, foundational, predictive, and important concepts captured?
- Can some criteria be removed or replaced?
- What were the true causes of lack of reproducibility? Can they be quantified?
- How can the true minimum standards for PBPK modeling be identified and tested?
- Would development of similar criteria for other fields of scientific literature be valuable for identifying reproducibility?



Count of PBPK model papers in this set by publication year, with amounts meeting each criteria category. The proportion of PBPK model papers with potentially reproducible **methods** has decreased and the proportion with potentially reproducible **results** has stayed flat, despite guidance documents cited below being published in 2004, 2006, 2007, 2010, and 2012.

Discussion on increasing PBPK reproducibility:

- **Reviewers** could contribute to results reproducibility by requiring presentation of adequate data for model validation, especially if parameter values are being fit.
- Chemical safety assessors could require fulfillment of reproducibility **criteria** before adopting a PBPK model (at a minimum; for discussion of other important components outside the scope of this paper that could also be added to a checklist, see Clark et al[6]).
- A **database of peer-validated code** would allow sensitivity analysis of parameters and systematic assessment of portability among models, as well as comparison of variability to aid in quantifying the inherent uncertainty in this type of work.
- Use of **open PBPK software** circumvents commonly observed reproducibility obstacles. For example, the openly available R package *httk* creates generic PBPK model results for hundreds of chemicals[7]. Differences between basic hepatic clearance model results and a given set of experimental data could identify when additional mechanisms are needed to adequately model a dose response.

5. References

(1) National Research Council, Committee on Toxicity Testing and Assessment of Environmental Agents, Board of Environmental Studies and Toxicology, Institute for Laboratory Animal Research, Toxicity Testing in the 21st Century. A Vision and a Strategy , 2007a Washington, DC National Academies Press **(2)** U.S. Environmental Protection Agency (EPA). (2006) Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment. National Center for Environmental Assessment, Washington, DC; EPA/600/R-05/043F. **(3)** McLanahan, E. D., El-Masri, H. A., Sweeney, L. M., Kopylev, L. Y., Clewell, H. J., Wambaugh, J. F., Schlosser, P. M. (2012). Physiologically Based Pharmacokinetic Model Use in Risk Assessment - Why Being Published Is Not Enough, *Toxicological Sciences*, 126(1), 5-15. **(4)** Barton, H. A., Chiu, W. A., Setzer, R. W., Andersen, M. E., Bailer, A. J., Bois, F. Y., DeWoskin, R. S., Hays, S., Johanson, G., Jones, N., Loizou, G., MacPhail, R. C., Portier, C. J., Spendiff, M. Tan, Y. (2007) Characterizing Uncertainty and Variability in Physiologically Based Pharmacokinetic Models: State of the Science and Needs for Research and Implementation, *Toxicological Sciences*, 99 (2), 395-402. **(5)** International Programme on Chemical Safety. Characterization and application of physiologically based pharmacokinetic models in risk assessment. (IPCS harmonization project document; no. 9). WHO 2010. **(6)** Clark, L. H., Setzer, R. W., & Barton, H. A. (2004). Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. Risk Analysis, 24(6), 1697–1717. **(7)** Pearce, R., Setzer, R., Strope, C., Sipes, N., & Wambaugh, J. (2017). httk: R Package for High-Throughput Toxicokinetics. *Journal of Statistical Software, 79*(4), 1-26. **(8)** MacLeod, M., Scheringer, M., McKone, T. E., & Hungerbuhler, K. (2010). The State of Multimedia Mass-Balance Modeling in Environmental Science and Decision-Making. *Environmental Science & Technology 44* (22), 8360-8364 | Figures implemented in Python 3.5 using matplotlib_venn | Analysis conducted using Python 3.5, R 3.3.2, MATLAB R2015a