Measuring Physicochemical Properties to Inform the Scope of Existing QSAR/QSPR Models

Chantel I. Nicolas[1,2,5], Kamel Mansouri[1,2,5], Katherine A. Phillips[3], Christopher M. Grulke[2], Ann M. Richard[2] , Andrew D. McEachran[2,5] , Antony J. Williams[2], James Rabinowitz[2], Kristin K. Isaacs[3], Alice Yau[4], and John F. Wambaugh[2]*

[1]Scitovation, Durham, NC; [2]National Center for Computational Toxicology and [3]National Exposure Research Laboratory , US Environmental Protection Agency, RTP, NC; [4]Southwest Research Institute, San Antonio, TX; [5]Oak Ridge Institute for Science and Education, Oak Ridge, TN

INTRODUCTION

Chemical structures and their properties are important for determining their potential toxicological effects, toxicokinetics, and route(s) of exposure. These data are needed to prioritize thousands of environmental chemicals, but are often lacking. In order to fill data gaps, robust quantitative structure-activity relationship (QSAR) and quantitative-structure property relationship (QSPR) models are routinely used in risk assessment for both well-known and new chemicals. However, all QSAR and QSPR models are limited in part by the training-set of data available for model development. In order to both calibrate and inform the scope of currently available QSPR models, physicochemical measurements were attempted for 200 chemicals selected for a mix of both those with previously measured physicochemical properties as well as chemicals with moieties that were expected to be challenging to existing models from the U.S. EPA DSSTox database. Among the properties measured were octanol:water partitioning coefficient (Kow), vapor pressure (VP), water solubility (WS), Henry's law constant (HLC), and acid dissociation constant (pKa).

OBJECTIVES

The **purpose** of the present work was to:

A. Determine an optimal chemical dataset to be submitted for measurements.

B. Compare these measurements with

1) previous measurements, if available, and

2) predicted values from various models.

C. Determine the impact of new measurements on the models that are sensitive to them (i.e., volume of distribution (V_D)).

D. Gauge the reproducibility of physicochemical property measurement methods.

E. Fill data gaps in measurement data (i.e., pKa).

F. Calibrate predictive models.

METHODS AND RESULTS

Test Chemical Selection

Figure 1 illustrates a structured workflow that was designed for the purpose of filtering chemicals in the DSSTox database.

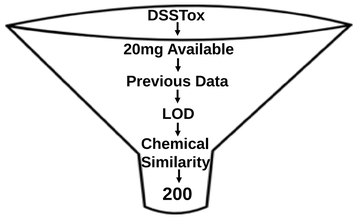


Figure 1. Flowchart describing analysis pipeline for picking chemicals.

In order to prepare a uniform list of 200 chemicals for each property measurement, the final test chemical set consisted of: ~ 20% with previous K_{OW}, HLC, and WS measurements and ~ 15% with previous VP measurements. The numbers of chemicals successfully measured for each property were: 176 (Kow), 168 (VP), 129 (WS), 110 (HLC), and 100 (pKa).

Structural Signatures Indicative of Failure

Figure 2 illustrates the structural similarity between chemicals for which K_{OW} measurements were successful (176) and those that were unsuccessful (23) using ToxPrint chemotypes from USEPA DSSTox. Out of 595 structural signatures, 66 signatures represented the chemical space of the 23 failed compounds. Only 25 signatures had a sum of greater than one (i.e., at least 2 or more compounds contained the same structural component. Major differences in representation of certain signatures may serve as a qualitative method for gauging trends in measurement efficacy based on structural features.

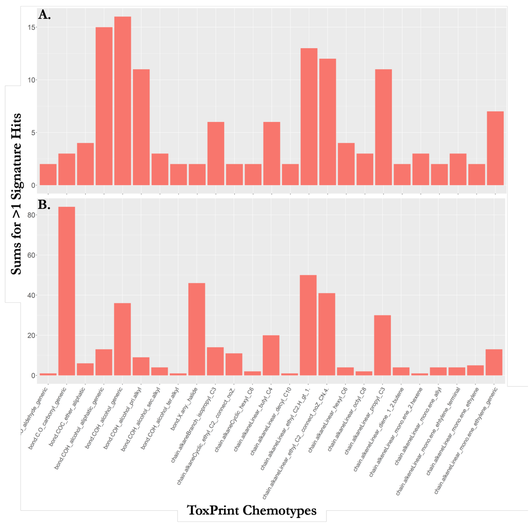


Figure 2. Comparative histograms for the sum of structural signatures: A. 23 K_{OW} measurement failures and B. 176 successful measurements.

K_{OW} Measurements vs Previous Measurements

As illustrated in Figure 3, an R2 of 0.77 is achieved when the newly measured Kow values are linearly regressed on previously measured values reported in EPI Suite. 176 out of 200 chemicals were returned with experimental values, while all 200 have previously measured values in EPI Suite (compiled in the PhysProp database). As red dotted lines indicate LOD for methods utilized, some measurements were still reported outside of the range: 0 < Kow < 6.

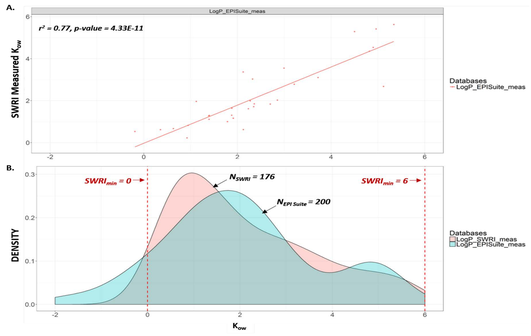


Figure 3. Comparative histograms for measured K_{OW}: A. Plot of new measurements (SWRI) vs EPI Suite measurements and B. Comparison of ranges for measured values (dotted lines indicate limits of detection).

K_{OW} Measurements vs Various Predictions

As illustrated in Figure 4, relatively similar R2 values are achieved (~0.6) when SWRI Kow measurements are regressed on predictions from ACD Labs, EPI Suite, NICEATM, and OPERA databases. While 176 chemicals were returned with the new measurements, there were predictions for nearly or all 200 chemicals from all databases.

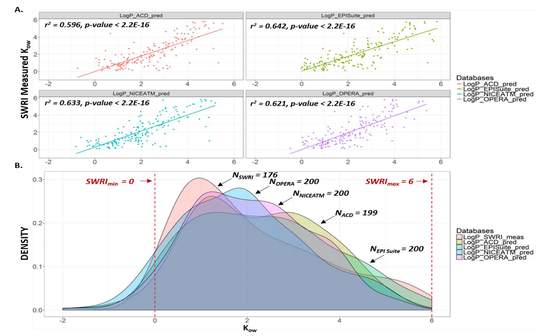


Figure 4. Comparative histograms for K_{OW}: A. Plot of new measurements (SWRI) vs EPI Suite measurements and B. Comparison of ranges for predicted values (dotted lines indicate limits of detection).

Physicochemical properties are among the various parameters necessary for describing the pharmacokinetics (PK) of chemicals of interest such as the volume of distribution (V_D). V_D is the apparent volume that would be needed to contain the observed plasma concentration of an administered compound. K_{OW} is used to predict V_D and they are shown to be linearly correlated. Figure 5 is a comparison of measured versus predicted values of volume of distribution.

Volume of Distribution

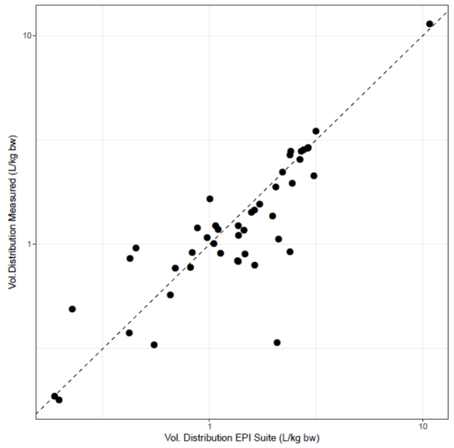


Figure 5. Impact of new K_{OW} measurements of V_D.

CONCLUSION

While one major aspect of obtaining new measurements for important physicochemical properties is the opportunity to fill data gaps, there are many other benefits. There are opportunities to calibrate existing PK models such as for V_D as well as to determine whether new measurement techniques ought to be developed for chemicals belonging to a distinct space that consistently fail.

REFERENCES

Mansouri, K., et al. (2016). "CERAPP: Collaborative Estrogen Receptor Activity Prediction Project." Environ Health Perspect. 124(7):1023-33

McEachran, A. D., et al. (2016). "Identifying known unknowns using the US EPA's CompTox Chemistry Dashboard." Analytical and Bioanalytical Chemistry: 1-7.

Wambaugh, J. F., et al. (2015). "Toxicokinetic Triage for Environmental Chemicals." Toxicol Sci 147(1): 55-67.

Richard, A. M. and Williams, C. R. (2002). "Distributed structure-searchable toxicity (DSSTox) public database network: a proposal." Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 499(1): 27-52.

USEPA (2015). EPI (Estimation Programs Interface) Suite.

Wetmore, B. A., et al. (2015). "Incorporating High-Throughput Exposure Predictions with Dosimetry-Adjusted In vitro Bioactivity to Inform Chemical Toxicity Testing." Toxicol Sci 148(1): 121-136.

ACKNOWLEDGEMENTS

Funding for this research was provided by the American Chemistry Council Long-range Research Initiative (ACC-LRI). The U.S. Environmental Protection Agency has provided administrative review and has approved this paper for publication. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. Environmental Protection Agency.

SciToVation
INNOVATIVE CELL BASED SCIENCE