Methodology for Uncertainty Analysis of Dynamic Computational Toxicology Models

Jimena Davis[1], John Wambaugh[1], Ramon I. Garcia[2], R. Woodrow Setzer[1]

U.S EPA, ORD, [1]National Center for Computational Toxicology; [2]UNC Chapel Hill, Dept. of Biostatistics

research&development

Science Question

With the increased use of complex computational toxicological models by the EPA for regulatory purposes, quantifying the uncertainty in model predictions is an important challenge that is being addressed with the following science questions:

- How can informative priors be developed to account for uncertainty in chemical-specific parameters in the absence of data?
- How can the issue of model identifiability be addressed in the analysis of physiologically-based pharmacokinetic (PBPK) model parameters in a Bayesian framework?
- How can computational issues (i.e., long runtimes) be addressed for hierarchical Bayesian analysis of PBPK models?
- What measures can be used to evaluate how well models describe data and model uncertainty?

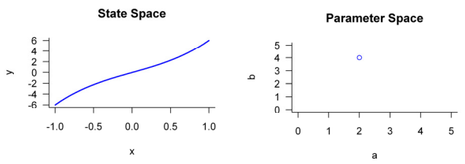
Research Goals

- Develop standard approaches for the specification of informative priors for chemical-specific parameters.
- Determine identifiability of model parameters and establish ways to handle unidentifiable parameters in Bayesian estimation.
- Develop efficient computational methods for model calibration and uncertainty analysis in a Bayesian setting.
- Adapt standard techniques for model evaluation of dynamic computational toxicology models and evaluate their efficacy in improperly and properly specified examples.

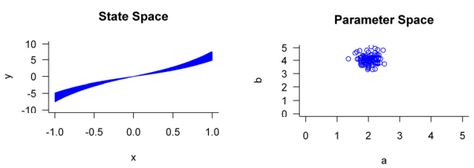
Methods/Approach

What is Uncertainty?

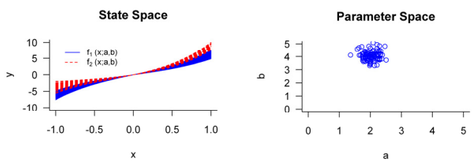
When developing computational models for various applications (e.g., physical, chemical, biological), the different types of uncertainty involved with the use of these models must also be considered. Two major types of uncertainty that should be addressed and quantified are ***parameter uncertainty and model uncertainty***.



Of the two forms, ***parameter uncertainty*** is the easiest to quantify via confidence intervals (standard errors) or probability distributions.



However, there is often a great deal of uncertainty associated with the model structure or formulation. Along with quantifying uncertainty in parameter estimates, model evaluation and selection are important with regards to addressing ***model uncertainty***.



Our goal is to develop a standard set of practices that will allow one to quantify both parameter and model uncertainty in a probabilistic framework efficiently.

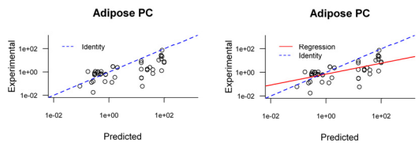
Quantifying Prior Knowledge for Chemical-Specific Parameters

- Uncertainties in model parameters and predictions for computational toxicology applications (e.g., ToxCastTM/high throughput situations, virtual tissues projects) can be quantified via prior information obtained from *in vitro* assays in the absence of *in vivo* data.

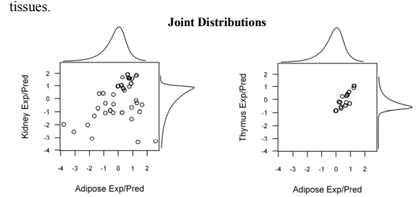
- More informative prior distributions (means and standard deviations) for Bayesian analysis of deterministic (e.g., PBPK) models can be developed with the use of computational predictors (e.g., QSAR and QSPR models), *in vitro* methods, and experimentally measured values for chemical-specific parameters found in the literature.

Example: Partition Coefficients

Partition coefficients (PCs) are important PBPK model parameters because of their role in determining the distribution of a compound throughout the body.

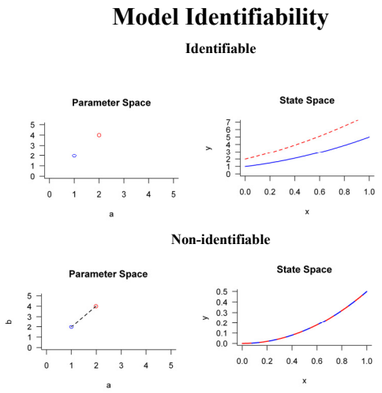


Coefficients of variation (CVs) were calculated mostly in the range of 50-70% for tissue:plasma PCs using Schmitt's (2008) computational predictor and experimental PCs data from the literature for various tissues.



➢ Milestone: Manuscript on constructing priors for PBPK model parameters (2009)

Model Identifiability

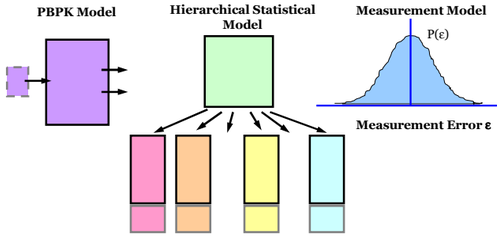


Garcia et al. (in preparation) observe that identifiability of PBPK model parameters can not be determined. However, specifying truncated prior distributions for parameters result in proper posteriors that are necessary for valid statistical analysis of PBPK models. ***Therefore, inferences based on Bayesian analyses that result in proper posteriors are valid even though the PBPK model may not be identifiable.***

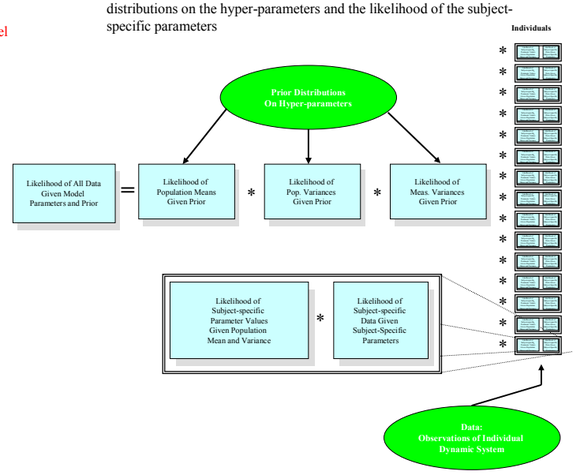
Hierarchical Bayesian Analysis for PBPK Models

Calibration of many PBPK models involves data from multiple experimental designs, endpoints, and laboratories; hence uncertainty and variability must both be addressed correctly in a statistical framework. Hierarchical statistical models lend themselves easily to the usage of multiple data sets and evaluate both uncertainty and variability for a large number of parameters.

Three components must be formulated for Bayesian analysis of PBPK models:



The likelihood of a set of data is given in terms of the prior distributions on the hyper-parameters and the likelihood of the subject-specific parameters



Advantages in using hierarchical Bayesian analysis in calibration and uncertainty analysis of PBPK models:

- Allows prior knowledge about parameters to be incorporated easily through priors assigned to parameters.
- Characterizes identifiability of parameters via comparison of priors with posteriors.
- Quantifies uncertainty and variability in model parameters through the estimation of distributions versus point estimates.

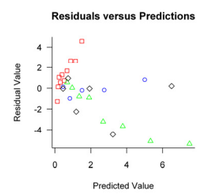
Computational Issues with PBPK Models

- One major drawback to the use of Bayesian analysis for PBPK models is the long computational runtimes often required when these models are implemented via Markov Chain Monte Carlo (MCMC) methods.
- The high dimensionality of the parameter space can cause computations to take weeks or even months. Convergence of MCMC runs to the posterior is also an issue that makes hierarchical Bayesian analyses very costly.
- Computations can be simplified via parallelization of the likelihood computations when using the standard model for hierarchical Bayesian analysis as demonstrated by Wambaugh et al. (in preparation)… hence decreased runtimes.

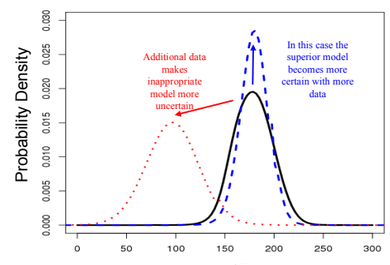
➢ Milestone: Manuscript on improvement of computational efficiency in Bayesian analyses of PBPK models using MCMC and assessment of convergence (2009)

Model Evaluation

- Model uncertainty can be addressed with Bayesian analysis by evaluating how well a model fits data as well as quantitatively comparing alternative model structures.
- Standard statistical approaches (e.g., residual plots) that exist for simpler linear regression models must be adapted for analysis of more complex toxicological models.



- We will develop examples using well-specified and ill-specified PBPK models and data to illustrate some of the problems that can arise and be addressed when evaluating computational toxicology models.



➢ Milestone: Manuscript on assessing the fit of PBPK models (and, through example, other complex mechanistic models) to data (2010)

Results/Conclusions

- Quantifying parameter uncertainty can be improved through the use of computational predictors and readily available data in the literature.

- Since it is not generally possible to assess the identifiability of parameters in PBPK models with a given data set, **only Bayesian analyses with proper priors** can be used for valid statistical inferences of PBPK models.

- Computational runtimes with MCMC methods for the generic model for hierarchical Bayesian analysis can be decreased by parallelization at the individual (or subject specific) level.

Impact and Outcomes

- Uncertainty quantification in coupled exposure-dose model for permethrin (see Poster #15). (Milestone: Science Advisory Panel (SAP) review in July 2010)

- Improved uncertainty analysis in cumulative risk assessment for pyrethroid pesticides in collaboration with other ORD laboratories (NERL and NHEERL) (Milestone: SAP review in 2012)

Future Directions

- Continued development of tools to address the computational issues associated with Bayesian analyses of PBPK models.

- Formulation of guidelines that can be used in model evaluation and selection for more complex computational toxicological applications.

- Identification and assessment of computational issues associated with model evaluation and selection for agent-based virtual tissues models.

References

Garcia, R.I., et al. (*in preparation)* Identifiability of PBPK models with applications to dimethylarsinic exposure.

Schmitt, W. (2008) General approach for the calculation of tissue to plasma partition coefficients, *Toxicology in Vitro*, **22**, 457 – 467.

Wambaugh, J.F., et al. (*in preparation*) Efficient Markov Chain Monte Carlo analysis of physiologically-based pharmacokinetic models.