Citation: CPT Pharmacometrics Syst. Pharmacol. (2014) 3, e101; doi:10.1038/psp.2013.77

www.nature.com/psp

PERSPECTIVE

Evaluating Systems Pharmacology Models Is Different From Evaluating Standard Pharmacokinetic–Pharmacodynamic Models

B Agoram[1]

Based on the author's recent experience, there appears to be some confusion regarding the steps required to qualify a systems pharmacology model as adequate for the intended purpose. This manuscript outlines the model evaluation approach used in the author's recent publication[1] on the systems pharmacology of a 5-lipoxygenase inhibitor and is an attempt to generate discussion on this topic within the pharmacometrics and systems pharmacology community.

CPT Pharmacometrics Syst. Pharmacol. (2014) **3**, e101; doi:10.1038/psp.2013.77; published online 19 February 2014

As the pharmacokinetic–pharmacodynamic (PKPD) modeling community realizes the value of systems pharmacology modeling in drug discovery and development and practices more of it, an understanding of some of the differences in the practice of modeling between the two communities is necessary. Model evaluation, for standard PKPD models, typically consists of two main steps:

1. Evaluation of the adequacy of the parsimonious model—i.e., a model with the fewest number of estimated parameters—to describe the data: this is typically done using goodness-of-fit plots of predicted and observed concentrations, and the weighted residuals, simulation-based checks such as visual predictive check, and comparison of the minimum objective function value.
2. Evaluation of the predictive capability of the model: though not as commonly applied as step 1, an external predictive check may also be sometimes applied to evaluate the ability of the model to predict outcomes which have not been used to train the model, if such a data set exists.

These model evaluation steps are often considered sufficient qualification for the typical uses of a PKPD model: selection of doses in a trial, description and interpretation of longitudinal study data, etc.

These steps may not always be possible, relevant, or adequate evaluation of a systems pharmacology model for various reasons. For example, in many cases, these models are used to generate hypotheses and are, therefore, exploratory in nature. Any imprecision or bias in the description of data may not be sufficient to reject the models as not useful. Indeed, as pointed out recently,[2,3] discrepancy between a systems model and observed data may itself be informative about the existing knowledge of an underlying mechanism. Unlike PKPD models, whose data come from controlled clinical and preclinical trials, the data for developing systems models tend to be drawn from various sources, including clinical, preclinical, and *in vitro* experiments, and thus may be associated with unknown uncertainty. Also, the most parsimonious model is not always the most useful model for developing a mechanistic understanding of the system.

This discrepancy in the approach to evaluating systems models vs. standard PKPD models has caused some confusion within the systems pharmacology and pharmacometrics community, an example of which we encountered during the review process of our manuscript published recently in this journal.[1] In that paper, a systems model of the arachidonic acid pathway was developed and used to understand the therapeutic value of blocking the 5-lipoxygenase (5-LO) enzyme. The following steps were taken to evaluate the model reported in that manuscript. In this author's opinion, these steps provide a convenient framework for evaluating systems models in general and are provided here to generate further discussion on this topic.

CLEAR STATEMENT OF THE OBJECTIVE OF THE MODELING EXERCISE

All model development and evaluation should be fit-for-purpose. Cliched as this sounds, it is worth remembering this point, since systems models are typically developed for objectives quite different from those for standard PKPD models. The objectives would determine the model scope—prior knowledge the model is built on, understanding of the physiology, the biological scale of the problem, the source of the model parameters, etc. In the study of Demin *et al.*,[1] the main objectives of the effort were to understand the mechanism behind the two-phase bronchodilatory response of zileuton, a 5-LO inhibitor, and to compare the relative bronchodilatory efficacy of 5-LO inhibition and leukotriene receptor blockade. Sufficient literature exists on the mechanism of action of 5-LO to formulate the system of equations; however, accurate human *in vivo* parameter values are not known. Therefore, the mathematical model structure, with parameters estimated from *in vitro* and *ex vivo* human and animal data, was considered sufficient for these objectives, in spite of their

[1]MedImmune LLC, Clinical Pharmacology, Drug Metabolism, and Pharmacokinetics, Cambridge, UK. Correspondence: B Agoram (agoramb@medimmune.com)
Received 16 September 2013; accepted 3 December 2013; advance online publication 19 February 2014. doi:10.1038/psp.2013.77

unknown uncertainty. Thus, model calibration and evaluation of predictive capability were not performed. For a more quantitative objective (e.g., study design), more accurate parameter values would be required and consequently, a robust data set would be necessary.[4] It is important that, along with stating the objective of the modeling exercise, the model evaluation procedure is also justified up-front.

JUSTIFICATION OF THE MODEL STRUCTURE

For many PKPD models, the structure is usually well accepted within the modeling community, and justification of the final model structure only requires support of the parsimony principle—i.e., that the model provides a sufficient goodness-of-fit of the data using the fewest number of parameters. Selection of a systems model structure is usually based on different considerations—on the basis of prior knowledge of the underlying mechanism only,[1,5] available data and/or prior knowledge of the mechanism,[4] or a desire to explore the ability of unknown mechanisms to describe data.[6] Prior knowledge of the underlying pharmacology is, therefore, critical to the justification of the model structure. Simple representations of the model that capture the essence of the underlying pharmacology are critical to justify the model to reviewers, especially when the model is complex or when the model structure is being reported for the first time. Another practice that greatly helps facilitate this interaction is if the full model code and all parameter values were made available to the reviewers and readers in a software-independent format (e.g., Systems Biology Markup Language) while submitting the manuscript for review, as is indeed required by this journal. For cases where exploratory models are evaluated, the choice of alternative models should also be clearly justified, since the number of alternative models can be large.

EVALUATION OF COMPONENT SUBMODELS

In many cases, the overall model is composed of individual subcomponents each describing different pharmacological interactions integrated either within the same scale[1] or across multiple scales.[5,7,8] The component submodels can be relatively simple (e.g., E_{max} type models) or more complex. Standard model development criteria—goodness-of-fit plots, visual predictive check, etc—can be applied to the evaluation of simple models, as is done by Demin *et al.*[1] Individual parameter sensitivity analyses can also be used to evaluate the impact of the uncertainty in the assumed parameter value on the emergent properties of the submodel.

QUALIFICATION OF THE EMERGENT PROPERTIES OF THE SYSTEM

The final systems model is typically used in one or more of the following ways: (i) to explore experimentally untested input–output response in a system; (ii) to understand complex behavior of a system; and (iii) to develop a quantitative understanding of all markers in a system. Objectives (i) and (ii) are semi-quantitative, therefore, previously listed model evaluation steps are considered sufficient to qualify the model as fit for these purposes even if the predictions and observations do not entirely agree. In our publication,[1] this was the case, the predicted maximum bronchodilation for zileuton was ~20–25%, whereas the reported value is ~15%. No effort was made to further refine model parameters to ensure a more accurate prediction. If an accurate and unbiased description of the data is required, further refinement of model structure and parameter values may be required. If so, the choice of parameter to be reestimated and the modifications to the model structure should be justified.[4]

DISCUSSION AND CONCLUSION

Recently, Hendriks[2] initiated a discussion on some pitfalls in the evaluation of systems models by methods commonly applied to empirical PKPD models. The intent of this commentary is to contribute to this emerging debate by illustrating how a systems model of 5-LO inhibitor was evaluated,[1] based on other recent systems pharmacology publications.[4,5] Possibly, this could provide a first framework for evaluating systems pharmacology models.

Empirical PKPD models are often confirmatory; therefore, a precise unbiased description of data is necessary to achieve this. For example, to confirm dose adjustment in a subpopulation, an accurate and unbiased description of the variability in exposure is required. However, a systems model may be useful even while providing an inaccurate description of data because such an inaccurate description could provide insights into false assumptions regarding the underlying pharmacology. Therefore, standard diagnostics (parsimony, goodness-of-fit plots, objective function value, visual predictive check, etc) should not be the sole criteria for evaluating systems models, even though correspondence between prediction and observation is necessary to engender confidence in the model. This dichotomy between subjective and objective criteria is at the root of the confusion in how to evaluate these models. In the author's experience, a clear statement of the objectives of modeling exercise can go a long way in both identifying and helping to justify the evaluation procedure.

If objective criteria such as the ones commonly used are not applicable, then what? Justification of a systems model on the basis of plausible structures and parameter values based on prior mechanistic knowledge and qualitative comparison of prediction vs. observations all constitute subjective criteria and should be acceptable for systems model justification. For example, in the study by Demin *et al.*,[1] the predicted bronchodilation for zileuton was 20–25% vs. ~15% observed. However, the conclusions were still considered valid because of a weight-of-evidence approach: accurate prediction of biomarker and bronchodilation changes with montelukast, the known assumptions in the model, and the qualitative rather than quantitative objective of the effort. Because of these limitations, it is prudent to often consider a systems model as *a plausible* and *useful* mathematical description of a system and not as *the definitive* description of the system.

Hopefully, further discussion of this topic will result in a more formal guidance for the modeling community. In the meanwhile, the set of points to consider in this commentary

may help provide a starting point for obtaining convergence between authors and reviewers of systems models and help map out a way forward for authors and peer reviewers to reach a compromise in evaluating these models.

Acknowledgments. The author acknowledges input from Lorin Roskos, and Mark Penney on draft versions of this document.

Conflict of Interest. The author declared no conflicts of interest

1. Demin, O. *et al.* Systems pharmacology models can be used to understand complex pharmacokinetic-pharmacodynamic behavior: an example using 5-lipoxygenase inhibitors. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e74 (2013).
2. Hendriks, B. Negative modeling results: a dime a dozen or a stepping stone to scientific discovery? *CPT Pharmacometrics Syst. Pharmacol.* **2**, e48 (2013).
3. Brännmark, C. *et al.* Insulin signaling in type 2 diabetes: experimental and modeling analyses reveal mechanisms of insulin resistance in human adipocytes. *J. Biol. Chem.* **288**, 9867–9880 (2013).
4. Landersdorfer, C.B., He, Y.L. & Jusko, W.J. Mechanism-based population modelling of the effects of vildagliptin on GLP-1, glucose and insulin in patients with type 2 diabetes. *Br. J. Clin. Pharmacol.* **73**, 373–390 (2012).
5. Peterson, M.C. & Riggs, M.M. Predicting nonlinear changes in bone mineral density over time using a multiscale systems pharmacology model. *CPT Pharmacometrics Syst. Pharmacol.* **1**, e14 (2012).
6. Toghaw, P., Matone, A., Lenbury, Y. & De Gaetano, A. Bariatric surgery and T2DM improvement mechanisms: a mathematical model. *Theor. Biol. Med. Model.* **9**, 16 (2012).
7. Cedersund, G. & Strålfors, P. Putting the pieces together in diabetes research: towards a hierarchical model of whole-body glucose homeostasis. *Eur. J. Pharm. Sci.* **36**, 91–104 (2009).
8. Nyman, E. *et al.* A hierarchical whole-body modeling approach elucidates the link between *in Vitro* insulin signaling and *in Vivo* glucose homeostasis. *J. Biol. Chem.* **286**, 26028–26041 (2011).