TUTORIAL

Sobol Sensitivity Analysis: A Tool to Guide the Development and Evaluation of Systems Pharmacology Models

X-Y Zhang*, MN Trame, LJ Lesko and S Schmidt*

A systems pharmacology model typically integrates pharmacokinetic, biochemical network, and systems biology concepts into a unifying approach. It typically consists of a large number of parameters and reaction species that are interlinked based upon the underlying (patho)physiology and the mechanism of drug action. The more complex these models are, the greater the challenge of reliably identifying and estimating respective model parameters. Global sensitivity analysis provides an innovative tool that can meet this challenge.

CPT Pharmacometrics Syst. Pharmacol. (2015) 4, 69-79; doi:10.1002/psp4.6; published online 25 February 2015

The need of qualitatively and quantitatively understanding the dynamics of complex biological systems has given rise to systems pharmacology models. The mechanistic nature of these models allows for a detailed description of the underlying network of biological processes and how they respond to therapeutic interventions. At the same time, the relative complexity of these models results in a number of challenges, one of which is the ability to identify and estimate respective model parameters. Models that are less complex but are still able to characterize the dynamic properties of underlying biological systems have consequently been gaining popularity. In order to develop these models, a number of approaches can be used, one of which is sensitivity analysis. Sensitivity analysis is routinely employed to evaluate how changes in model input affect its output and, thus, to determine to what extent changes in a model input parameter or a set of model input parameters will affect the model output.

In this tutorial, we will first introduce the concept of sensitivity analysis before discussing the two general subtypes of sensitivity analysis, i.e., local and global. Following this general introduction, we will briefly discuss the advantages and limitations of the different approaches used for global sensitivity analysis. Finally, we will discuss Sobol's method, the associated analysis steps as well as its advantages and limitations. We will follow with three specific examples in an increasing order of complexity.

SYSTEMS PHARMACOLOGY MODELS

Quantitative and systems pharmacology has been regarded as an important tool and computational approach to discover novel drugs and elucidate the mechanism of drug actions. It has been well accepted that biological systems consist of highly interlinked networks, and the dynamic interplay between the individual components needs to be better understood in order to make inferences about the overall system's behavior. Without a thorough understanding of the intricacies

of biological systems, it is difficult to understand and predict the mechanisms of drug action at the molecular level. Mathematical modeling and simulation tools can help to overcome this challenge as they can be used to integrate information from different *in vitro*, animal, and clinical studies into a single, unifying model.³ Respective models can be established at different levels of spatial and temporal complexity ranging from data-driven and descriptive to completely mechanistic approaches as in systems pharmacology modeling.⁴

Models characterizing biological systems and drug actions can be established at different levels of spatial and temporal complexity using either a top-down or a bottom-up approach. Top-down approaches usually start at a high level of organization and become increasingly more complex to better characterize the underlying biological system. Bottom-up approaches, on the other hand, start with the "bottom" elements of the organism (e.g., genes or proteins and their known interactions), and work their way up to the patientphenotype level. 5,6 Both approaches have advantages and limitations, which have given rise to hybrid approaches such as pharmacokinetic/enhanced-pharmacodynamic (PK/ePD) models, that represent a synergy between the two approaches. This hybrid approach merges system biology (i.e., bottom-up) with the traditional PK (i.e., top-down) models. A traditional PK/PD model characterizes and predicts drug dispositions and the effects of drugs in the living organisms under physiological and pathophysiological conditions, whereas the systems biology model describes the underlying biochemical network.7-9 This type of PK/ePD modeling has numerous advantages in incorporating the details of single or multiple drug concentrations, systems biology models, the pertinent regulatory networks and motifs of feedback/feed forward loops, individual genomic and epigenetic characteristics, which has made the individual patient therapy (precision medicine) viable. 10,11 Accordingly, a systems pharmacology model is a complex, dynamic model that consists of both drug-specific and biological system-specific components.

These components are reflective of the pharmacokinetic, biochemical network, and systems biology properties that contribute to the function of the overall system. These types of models have recently gained popularity in drug discovery and research, and have been applied in different therapeutic areas, particularly in oncology, diabetes, and pain. 12–14

Systems pharmacology models are set up to characterize the functioning of the overall system by a detailed description of the underlying physiological biochemical networks through the use of ordinary differential equations. 12 Given the highly complex nature of systems pharmacology models, the uncertainty in the model is related not only to a few parameters, e.g., bioavailability and absorption rate, as in a traditional PK model, but comes from two sources: (i) the structural model itself and (ii) its parameters. 15 This is at least in part the result of regulatory feedback and feed forward mechanisms that are included in systems pharmacology models, which allows cells to maintain homeostasis while responding to changes in the surrounding environment. 16,17 The parameters in a systems pharmacology model are often unknown, being not measurable, and having experimental variations in different cell lines and experimental conditions. The fact that a given set of data may be characterized by different input parameter combinations equally well does not mean that all of these parameters combinations are equally meaningful, e.g., from a physiological point of view. This poses a certain challenge, particularly for more complex models, when attempting to identify an optimal model parameterization. This uncertainty around which set of structural model parameters is most appropriate to characterize the data is frequently referred to as "sloppiness" in the systems pharmacology arena. 18-20

Given that uncertainty around the underlying structural model and identifiability of its parameters are the major challenges for the successful implementation of systems pharmacology models to problem-solving and decision-making, sensitivity analysis has lent itself as a viable solution and a reliable methodology that allows to understanding the relative importance of key processes and/or parameters driving the output dynamics of the overall system. Sensitivity analysis is also essential for guiding the further refinement of a systems pharmacology model and its application in drug discovery and development, e.g., for identifying drug targets or understanding the heterogeneity of drug response (efficacy/safety) differences.^{11,21–23}

Sensitivity analysis

Sensitivity analysis allows the identification of the parameter or set of parameters that have the greatest influence on the model output. It consequently provides useful insight into which model input contributes most to the variability of the model output.²⁴ Sensitivity analysis has been widely used in fields, such as risk assessment, economics, and engineering, and it has become instrumental in the systems pharmacology arena to guide the understanding and development of a complex model.^{25–27} The application of sensitivity analysis can be summarized as: (i) understanding the input—output relationship, (ii) determining to what extent uncertainty in structural model parameters contribute to the overall variability in the model output, (iii) identifying the

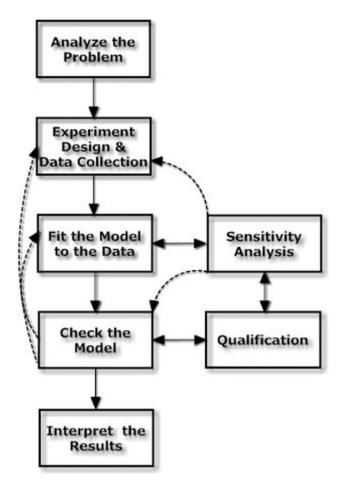


Figure 1 The flow chart and steps of data analysis, sensitivity analysis, the model establishment and qualification process. Sensitivity analysis can be used to guide the experiment design and throughout the model development process.

important and influential parameters that drive model outputs and magnitudes, and (iv) guiding future experimental designs. 24,28-30 For model builders and users, it is also a useful tool to check the model structure and uncertainty around the input parameters, and feedback into the model refinement to gain additional confidence in the model. Especially in a very complex model, the results of sensitivity analysis will help the model builders to focus on the critical parameters that determine the model output. **Figure 1** illustrates the steps of sensitivity analysis in guidance of an experimental design, parameter estimation during the model establishment, and qualification processes.

In general, there are two types of sensitivity analysis: (i) local and (ii) global.

Local sensitivity analysis

Local sensitivity analysis evaluates changes in the model outputs with respect to variations in a single parameter input. 31,32 The input parameters are typically changed one at a time in relatively small increments (e.g., 0.1% fold), 33 and the effect of this individual parameter perturbation on the model output is calculated using local sensitivity indices.

Table 1 Various types of commonly used global sensitivity analysis techniques comparison

Criteria for comparison	Commonly used global sensitivity analysis methods				
	Weighted average of local sensitivity analysis (WALS)	Partial rank correlation coefficient (PRCC)	Multi-parametric sensitivity analysis (MPSA)	Fourier amplitude sensitivity analysis (FAST)	Sobol
Discrete inputs	Yes	Yes	Yes	Yes	Yes
Model independence	No	No	No	Yes	Yes
Non-linear, input-output relationship	Yes	Yes	Yes	Yes	Yes
Non-monotonic input-output relationship	Yes	No	Yes	Yes	Yes
Robustness	Yes	Yes	Yes	Yes	Yes
Reproducibility	Yes	Yes	Yes	Yes	Yes
Ability to apportion the output variance	No	No	No	Yes	Yes
Higher order interaction of parameters	No	No	No	Yes	Yes
Quantitative measure for ranking	Yes	Yes	Yes	Yes	Yes
Computational efficiency	Yes	Yes	Yes	No	No

For example, consider a model given by a set of ordinary differential equations (y_i is the ith model output) that is dependent on a certain parameter set p and initial conditions $y_i(0)$:

$$\frac{dy_i}{dt} = f_i(y_i, p, t)i = 1, 2, ..., n.$$
 (1)

The local sensitivity indices, mathematically, being the first-order partial derivative of model outputs y_i respect to the model parameter p, can be calculated as: $\partial y_i/\partial p = \lim_{\Delta p \to 0} \frac{y_i(p+\Delta p)-y_i(p)}{\Delta p}$. Δp is the perturbation to the input parameter p, and it is usually a very small change of parameter p (e.g., 0.001^*p).

Local sensitivity analysis may only be used when the model output is linearly related to the parameters near a specific nominal value. The main limitation of a local sensitivity analysis is that it evaluates parameters one at a time, and does not allow for the evaluation of simultaneous changes in all model parameters. In addition, the interaction between parameters cannot be evaluated using a local sensitivity analysis. In order to overcome these limitations, global sensitivity analysis should be used.

Global sensitivity analysis

In a global sensitivity analysis, all parameters are varied simultaneously over the entire parameter space, which allows to simultaneously evaluate the relative contributions of each individual parameter as well as the interactions between parameters to the model output variance. Given that model inputs can span a wide range (e.g., the rate constants and initial concentrations) for systems pharmacology models, global sensitivity analysis is an innovative approach for determining which reactions and processes contribute most to the behavior of the overall system.

To date, several types of global sensitivity analyses, such as weighted average of local sensitivity analysis, partial rank correlation coefficient, multiparametric sensitivity analysis, Fourier amplitude sensitivity analysis (FAST) and Sobol's method, are available for application to systems pharmacology models.^{27,34–36} While an

in-depth discussion of these different techniques is beyond the scope of this tutorial, a brief description of these methodologies is provided below. In addition, their important features including advantages and limitations are summarized in **Table 1**.

The weighted average of local sensitivity analysis method calculates local sensitivity indices at multiple random points within the parameters space. The average of the local sensitivity indices is then used in combination with a weighting factor to approximate the global sensitivity.³⁷

The relationship between input and output is usually non-linear in complex systems pharmacology models. If the input—output relationship is nonlinear but still monotonic in a given model, a partial rank correlation coefficient can be used to identify the key parameters that determine the output using Pearson correlation coefficients. The values of partial rank correlation coefficient sensitivity indices range between -1 and 1. Zero values indicate that the model output is completely insensitive to changes in respective parameter inputs.³⁸

Latin hypercube sampling is a sampling method used to generate random parameter vectors. These randomly generated parameter vectors can then be used in simulations to map out the uncertainty in the model output that results from the uncertainty in the model parameter input during a multiparametric sensitivity analysis.³⁹ The vertical difference between acceptable and unacceptable parameters cumulative distribution is used to reflect the sensitivity of the multiparametric sensitivity analysis. The sensitivity values of a multiparametric sensitivity analysis range between 0 and 1. The closer the value is to 1, the more important the respective value is.^{39–41}

Both FAST and Sobol's methods are based on variance decomposition techniques to provide a quantitative measure of the contributions of the input to the output variance. The extended FAST (eFAST) is in extension of FAST technique, with the capability of calculating both total-order and first-order sensitivity indices. The main difference between FAST and Sobol's method is the underlying algorithm in multidimensional integration of the sensitivity indices. While the pattern search uses a sinusoidal function in FAST

method, a Monte Carlo integration method is used in Sobol sensitivity analysis. $^{\rm 42-45}$

Among all the presented global sensitivity analysis methods, variance decomposition based Sobol sensitivity analysis is so far one of the most powerful techniques (**Table 1**).

Sobol sensitivity analysis

Sobol's method is based on decomposition of the model output variance into summands of variances of the input parameters in increasing dimensionality. Sobol sensitivity analysis determines the contribution of each input parameter and their interactions to the overall model output variance.

Sobol sensitivity analysis is intended to determine how much of the variability in model output is dependent upon each of the input parameters, either upon a single parameter or upon an interaction between different parameters. The decomposition of the output variance in a Sobol sensitivity analysis employs the same principal as the classical analysis of variance in a factorial design. It should be noted that Sobol sensitivity analysis is not intended to identify the cause of the input variability. It just indicates what impact and to what extent it will have on model output. As a consequence, it cannot be used to determine the source(s) of variance, such as the impact of demographic covariates on total clearance, as in a conventional population PK/PD analysis.

One of the important steps in any sensitivity analysis, whether local or global, is to determine the appropriate model output to be used for the analysis. In theory, one could either use a certain point in time, such as C_{max} for a concentration-time curve, or a metric that integrates changes in the model output of interest over time, such as the duration of electrical depolarization, the integrated area under a drug plasma concentration-time curve, or even tumor size. The choice of the appropriate metric to be used depends on the question at hand. In most cases, the use of an integrated measure, such as the area under the concentration-time curve (AUC), provides a more robust metric of the model output. It should be noted at this point that an integrated metric can also change over time, for example, as the result of aging or disease (cf. disease systems analysis).

Sobol sensitivity analysis has several features listed as the following:

- · No assumption between model input and output
- Evaluation of the full range of each input parameter variation and interactions between parameters
- High computation intensity being the main drawback

There are important steps of Sobol sensitivity analysis, which will be discussed in greater detail in the following sections and are shown in **Figure 2**. First, a parameter sequence is generated using Sobol sequence. Sobol sequence, first proposed by Russian scientist I.M. Sobol, is a quasirandomized, low-discrepancy sequence. Low-discrepancy sequences typically sample space more uniformly than completely random sequences. Algorithms which use such sequences may have superior conver-

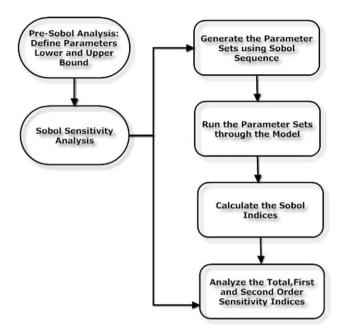


Figure 2 The flow chart and steps in implementation of a Sobol sensitivity analysis. There are two main steps: pre-Sobol and Sobol sensitivity analysis. The Sobol sensitivity analysis is divided into four steps: generating parameter sets, running and simulation the model output with the generated parameter sets, calculating, and analyzing the total-, first-, and second-order and higher-order Sobol sensitivity indices.

gence. 46 The generated parameter sets are subsequently used to simulate the model outputs. 47

The general features of Sobol sequence are listed as the following:

- Sobol sequence is a low-discrepancy sequence, also known as the "quasi-random sequence"
- More uniformly distributed than the pseudorandom numbers
- Quasi-Monte Carlo integration produces faster convergence and better accuracy
- · Drawback is the high-dimensional integrals

In order to understand how the output variance be attributed to individual input variables and the interaction between each of the input variables, the total-order, first-order, second-order, and higher-order sensitivity indices are calculated to accurately reflect the influence of the individual input, and the interaction between them.³⁶

Let x=(x1,x2,...,xs), be the input parameters. Each parameter is considered to range over some finite interval which may be assumed, after rescaling, to be [0, 1]. It is useful to think of each parameter as a random variable uniformly distributed on [0, 1], with all the parameters mutually independent. The model output whose sensitivity to the input parameters is to be assessed is a function of x, say f(x). Under the probabilistic interpretation of the parameters, f(x) is a random variable with mean (f_0) and variance (D):

$$f_0 = \int f(x) dx$$

$$D = \int f(x)^2 dx - f_0^2.$$

All integrals are multiple integrals with limits [0, 1] on each dimension.

The Sobol method is based on the decomposition of D into contributions from effects of single parameters, combined effects of pairs of parameters, and so on. This is done first by decomposing f(x) into

$$f(x) = f_0 + \sum_{i=1}^{s} f_i(x_i) + \sum_{i=1}^{s} \sum_{j \neq i}^{s} f_{ij}(x_i, x_j) + \dots f_{1 \dots s}(x_1, x_2, \dots, x_s).$$
 (2)

The terms of the decomposition are constructed as $follows^{36}$:

$$f_{i}(x_{i}) = \int f(x) \prod_{k \neq i} dx_{k} - f_{0}$$

$$f_{ij}(x_{i}, x_{j}) = \int f(x) \prod_{k \neq i, j} dx_{k} - f_{0} - f_{i}(x_{i}) - f_{j}(x_{j})$$

and so on.

The analysis of variance representation of f(x) is based on the satisfaction of the condition shown in Eq. 3^{36} :

$$\int f_{i1}, ..., i_{s}(x_{i1}, ..., x_{is}) dx_{k} = 0 \text{ for } k = i_{1}, ..., i_{s}.$$
(3)

Because of this property, squaring both sides of Eq. 2 and integrating yields

$$D = \sum_{i=1}^{k} D_i + \sum_{i < j} D_{ij} + \sum_{i < j < l} D_{ijl} + \dots + D_{1,2,\dots,k}$$
 (4)

where

$$D_{i_1...i_s} = \int f_{i_1...i_s}^2(x_{i_1},...,x_{i_s}) \, dx_{i_1},...,x_{i_s} \quad \text{is the variance of} \\ f_{i_1...i_s}(x_{i_1},...,x_{i_s}), \text{ called the partial variance corresponding to} \\ \text{that subset of parameters.}$$

The Sobol sensitivity indices for that subset of parameters is then defined as

$$S_{i_1...i_s} = \frac{D_{i_1...i_s}}{D}.$$
 (5)

For example, $S_i = \frac{D_i}{D}$ provides the first-order contribution from i_{th} input parameter to the output variance and $S_{ij} = \frac{D_{ij}}{D}$ is used to compute the second-order contribution from interaction between i_{th} and j_{th} parameters. Finally, total-order sensitivity indices, which are defined as the sum of all the sensitivity indices as $S_{Ti} = S_i + S_{ij:i \neq j} + \cdots + S_{1...i...s}$, quantify the overall effects of one parameter on the model output.

Respective sensitivity indices can be obtained from Eq. 4 by dividing it by D. Given that S_i relates the partial variance

to the total variance for each parameter, respective sensitivity indices should sum up to 1 as shown in Eq. 6.

$$1 = \sum_{i=1}^{k} S_i + \sum_{i < i} S_{ij} + \sum_{i < i < l} S_{ijl} + \dots + S_{1,2,\dots,k}.$$
 (6)

In summary, first-order sensitivity indices as the main effect are used to measure the fractional contribution of a single parameter to the output variance. Second-order sensitivity indices are used to measure the fractional contribution of parameter interactions to the output variance. Totalorder sensitivity indices take into account both the main, second-order and higher-order effects, which involves the evaluation over a full range of parameter space. The higher the sensitivity indices value, the more influential respective model parameters and the associated steps are. Although no distinct cutoff value has been defined, the rather arbitrary value of 0.05 is frequently accepted for this type of analysis for distinguishing important from unimportant parameters. It should be noted though that this value of 0.05 is primarily used for more complex models and it may be not stringent enough for relatively simple models that contain only few input parameters.

General steps of performing a Sobol sensitivity analysis

To date, the wide-spread application of this powerful analysis technique has been limited by the availability of open access software tools. However, the situation has improved with the emergence of public available software, such as the multiobjective evolutionary algorithm (MOEA) Framework (http://www.moeaframework.org/). The MOEA Framework is a library of multiple objective optimization evolutionary algorithms, with Sobol sensitivity analysis being one of the optimization libraries. During the sensitivity analysis procedure, MOEA Framework is treated as a black box. The optimization library is written in Java, and is thus platform independent. As a consequence, it is able to run on Mac, Linux, and Windows operating systems.

Prior to performing a Sobol sensitivity analysis, the right Java running environment needs to be set up in order to tailor to your computational operating systems. In general, there are two major steps in Sobol sensitivity analysis using the MOEA Framework (c.f. **Figure 2**): Step 1, pre-Sobol sensitivity analysis and Step 2, Sobol sensitivity analysis.

- · Pre-Sobol sensitivity analysis:
 - · Decide on the parameters in the model to be varied
 - Define the parameter range including the lower and upper bounds
 - Save the parameter values in a .txt file (e.g., "parameterul.txt")
- · Sobol sensitivity analysis:
 - Download the executable MOEA Framework (http://www.moeaframework.org/)
 - Generate the parameter sets using Sobol sequence with the following commands (in one line in the MOEA Framework):

java -classpath MOEAFramework-2.0-Executable.jar org.moeaframework.analysis.sensitivity.SampleGenerator -m saltelli -n 5000 -p parameterul.txt > sobolParameter.txt

Note:

- -m represents the mode of mathematical method of saltelli
 - -n represents the number of initial samples to generate from the pseudo-random Sobol sequence
 - -p represents the file specified in the pre-Sobol sensitivity analysis step

For example, if a model consists of 5 parameters (p), and 5,000 (n) number of initial sampling, then the total parameter sets will be generated using the equation 2n(p+1)=60,000

SobolParameter.txt is the parameter set that will be used in later steps for the model simulation.

Using the generated parameter sets to simulate the running model output

Note this step will be carried out in your running model;

 The outputs from your running model will be used to calculate the total/first/second-order sensitivity analysis in the MOEA Framework using the following command:

java -Xmx256m -classpath MOEAFramework-2.0-Executable.jar org.moeaframework.analysis.sensitivity.SobolAnalysis -m 0 -r 1000 –i objectiveValues.txt -p parameterlu.txt -o sobolIndices.txt

Note:

- -m represents the column of the output values (0 represents the first column)
- -r represents the number of bootstrap resamples
- -i represents the calculated output values from the previous step
- -p represents the file specified in the pre-Sobol sensitivity analysis step
 - -o represents the expected Sobol indices files
- Analyze the Sobol sensitivity indices and interpret the results:

Sobol sensitivity analysis is generally used for complex system models and it quantitatively decomposes the output variance with respect to its resources: i.e., from individual parameters or from the interaction between parameters. The total-order sensitivity indices are typically used to evaluate the overall contribution of a parameter and interaction with other parameters. The Sobol sensitivity indices have several features, which can be categorized as following:

The total-/first-/second-order sensitivity indices are positive values.

- Parameters with sensitivity indices greater than 0.05 are considered significant.
- The sum of all the sensitivity indices should be equal to 1.
- The total-order sensitivity indices are greater than the first-order sensitivity indices.

The sample size needed for performing a reliable Sobol sensitivity analysis depends on two main factors: (1) the complexity of the model and (2) the number of parameters evaluated. Although there is no general consensus on the optimal number of parameter sets to be generated, the general rule of thumb is that the larger numbers of model parameters, the higher the number of parameter sets to be used. For example, for a complex model with a large number of uncertain parameters (e.g., 20 parameters), at least 100,000 model evaluations should be performed. For less complex models, a smaller number of evaluations (e.g., 1,000) may be sufficient. It should be noted though that as the number of evaluations increases, the computational cost increases. The appropriateness of the number of evaluations selected can be tested using the bootstrap confidence intervals. Generally speaking, the most sensitive parameters should have narrow confidence intervals, which are less than 10% of the sensitivity indices.

Figure 3 shows the result of a hypothetical sensitivity analysis example of a sensitivity analysis result from an evaluation of 20 parameters in a systems pharmacology model containing 20 parameters. The results indicate that parameter k_{18} is the most important parameter contributing to ~25% of the model output variability, followed by the important parameters k_{17} , k_{16} , k_{10} , k_{8} , k_{5} , and k_4 . Both total-order sensitivity indices and first-order sensitivity indices have similar values, which indicate no significant second-order interaction between the parameters. The error bars in the figure represent the bootstrap confidence intervals (1.96*standard error) and 95% confidence intervals, which can be calculated by using the formula (sensitivity indices ± the bootstrap confidence intervals). If the confidence interval of the dominant parameter is greater than 10% of itself, the sample size needs to be increased by increasing the initial sampling number in the first parameter sets generation step.

Following a brief introduction of the global sensitivity analysis concept and the methodology used, we will now present three examples of increasing complexity starting with a simple PK case study in order to highlight the utility of the Sobol sensitivity analysis.

CASE EXAMPLES

Pharmacokinetic two-compartmental model

Our first example will use Sobol sensitivity to identify the key parameters that drive the plasma concentrations of the small molecule-Sunitinib, a tyrosine kinase inhibitor, which pharmacokinetics can be characterized by a two-compartment model with first-order absorption.⁴⁸

Following a single 50-mg oral dose of Sunitinib, the plasma concentrations were simulated over the course of 24 h. The

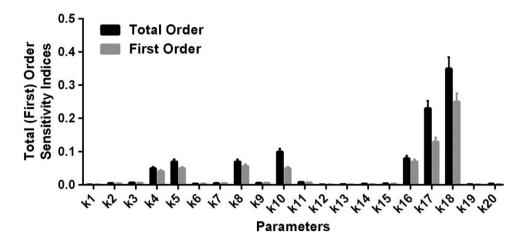


Figure 3 Sobol sensitivity indices of twenty kinetic parameters are shown in this graph. Total-order sensitivity indices (black bar) and first-order sensitivity indices (grey bar) are shown, respectively. The greater the sensitivity indices are, the more critical parameters are for the model. Complex models usually have more than one parameters that are critical and often vary at the same time with external conditions.

parameter values used in the Sobol sensitivity analysis varied between 100-fold lower and higher than the mean values obtained from a population model, except bioavailability, which varied between 0.5 and 1. 11,48 Ideally, parameter values that are constrained within the physiological lower and upper limits should be used in the pre-Sobol sensitivity analysis steps. Given that not all parameters of a systems pharmacology model translate one-to-one into a physiological effect, a rather arbitrary 100-fold range is used to test the impact of these parameters on model output.

Step 1: Parameters lower and upper bounds are saved in "sunitinib.txt" in the dataset files (pre-Sobol sensitivity analysis);

Step 2: Sobol sensitivity analysis procedures are listed in the followings and the files "sobolset.txt", "AUCPK.txt", and "sobolindices.txt" can be found in the attached dataset files.

- Set up the right Java running environment (e.g., for Windows XP operating systems, Cygwin can be downloaded from https://www.cgwin.com).
- Download the executable MOEA Framework 2.0 version (http://www.moeaframework.org/).
- The following commands are performed in the Cgwin interface in order to generate the parameters sets using the Sobol sequences:

java -classpath MOEAFramework-2.0-Executable.jar org.moeaframework.analysis.sensitivity.SampleGenerator -m saltelli

-n 50000 -p sunitinib.txt> sobolset.txt

 The file "sobolset.txt" is the generated parameter sets that will be used for simulation of the model output; "sobolset.txt" contains 2*50,000*(6+1) = 700,000 parameter sets.

Note: Even though no upper limit is set for parameter sets generation, we have adopted the relative large number

of parameter sets that will produce robust sensitivity indices with the acceptable computational efficiency.

 The generated parameter sets will be used for the model output simulations:

This step will be carried out in Matlab in simulation of the model output (integrated plasma concentration over 24 h (AUC)) values saved in "AUCPK.txt" file.

 Calculation of the total-/first-/second-order sensitivity indices:

java -Xmx256m -classpath MOEAFramework-2.0-Executable.jar org.moeaframework.analysis.sensitivity.SobolAnalysis -m 0 -r 1000 -i AUCPK.txt -p sunitinib.txt -o sobolindices.txt

Analyze and interpret the total-/first-/second-order sensitivity indices results saved in the "sobolindices.txt" as shown in Figure 4.

The results indicate that systemic clearance (cls), central compartment volume (vcs), and bioavailability (F) are three important parameters, in which vcs is the most dominant parameter with a total-order sensitivity index being 0.79 (Figure 4). For cls, if only the first-order sensitivity index is calculated, it is not as critical as when the total-order sensitivity was considered. This suggests that cls interacts strongly with other parameters. As expected, the secondorder sensitivity indices between parameters cls and vcs are 0.309, which indicates that the interaction between distribution and elimination (mainly in liver) is important for determining the variability of sunitinib plasma concentration within 24 h. The outcome of this relatively simple Sobol sensitivity analysis example shows that both clearance and volume of distribution are important parameters that drive the output of the model as reflected by the first-order sensitivity indices. The second-order sensitivity indices indicate that there is also a significant interaction between the two, which has to be considered in order to appropriately interpret the analysis output. While this insight has limited instructive value for the user of the Sobol sensitivity analysis for this simple case example, respective first- and second-order Sobol sensitivity indices values can provide very valuable insights into the underlying dynamics of highly complex systems as discussed in the following examples.

The pharmacokinetics of Naratriptan, a small molecule 5-HT₁ agonist, can be characterized by a two-compartment model. Variance decomposition was performed using the eFast method.⁴⁹ The sensitivity analysis shows that bioavailability is the dominant parameter that primarily governs

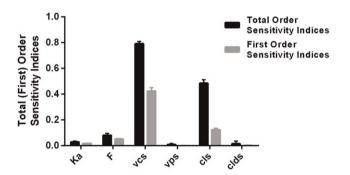


Figure 4 Sobol sensitivity indices of six kinetic parameters from a two-compartment Sunitinib pharmacokinetic model were shown in this graph. Total-order sensitivity indices (black bar) and first-order sensitivity indices (grey bar) were shown, respectively. vcs, cls, and F were critical parameters to the model output compared to other parameters. For cls, if only the first-order sensitivity indice was taken into consideration, it was not as critical as when the total-order sensitivity indice was considered. This suggested parameter cls interacted strongly with other parameters (vcs).

Naratriptan's pharmacokinetics, followed by volume of distribution of the central compartment and systemic clearance. Despite the fact that the pharmacokinetics of both Sunitinib and Naratriptan can be characterized by two-compartment body models, the Sobol sensitivity analysis using the MOEA framework (Sunitinib) is easier to implement and interpret compared to the eFast method (Naratriptan) previously published in accordingly, we encourage the simplifier yet powerful Sobol sensitivity analysis for model building and evaluation.

Pharmacodynamic HIV infection model

In our next example, we will use Sobol sensitivity analysis for a published human immunodeficiency virus (HIV) infection model in order to understand the key parameters that determine the free virus (V) model output. The results of this sensitivity analysis can be used to guide the therapeutic intervention strategy for eradicating free Virus V.

A mathematical model for HIV infection of CD4+ T cells was adopted from the literature. The structural model is composed of four compartments: uninfected T cells (Eq. 6), latently infected T cells (Eq. 8), actively infected T cells (Eq. 9) and free Virus (Eq. 10) (**Figure 5**). The model consists of four ordinary differential equations (Eqs. 6–10) and nine parameters, with the initial condition for T compartment (1,000 mm $^{-3}$), Virus HIV population V compartment (10 $^{-3}$ mm $^{-3}$), T* (0) and T** compartment (0), respectively. The structural model is comparable to the structural model is c

$$dT/dt = s - \mu_T T + rT(1 - (T + T^* + T^{**})/T_{\text{max}}) - k_1 VT$$
 (7)

$$dT^*/dt = k_1 VT - \mu_T T^* - k_2 T^*$$
 (8)

$$dT^{**}/dt = k_2 T^* - \mu_b T^{**}$$
 (9)

$$dV/dt = N_V \mu_b T^{**} - k_1 V T - \mu_V T. \tag{10}$$

The analysis steps outlined earlier in the paper were used to calculate the Sobol sensitivity indices. 100,000

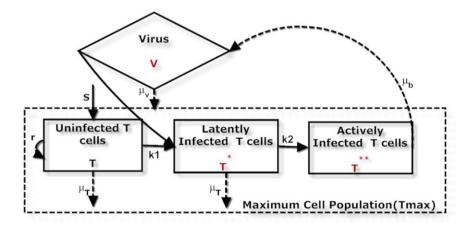


Figure 5 HIV infection diagram including free Virus (V), uninfected T cells (T), latently infected T cells (T*), and actively infected T cells (T*) four compartments. s (the rate of supply of CD4+ T cells from precursors), μ_T (the death rate of uninfected and latently infected CD4+ T cells), r (the rate of growth for the CD4+ T cell population), k_T (the rate constant for CD4+ T cells becoming infected by free virus), k_Z (rate latently infected cells convert to actively infected), μ_D (the death rate of actively infected CD4+ T cells), N_V (the number of free virus produced by lysing a CD4+ T cell), μ_V (the death rate of free virus), and Tmax (the maximum CD4+ T cell population) were adopted from the previous publication.⁵⁰

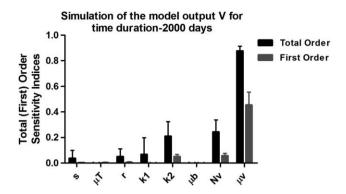


Figure 6 The total-order and first-order sensitivity indices for the output free virus V. Total-order (black bar) and first-order sensitivity indices (gray bar) were shown, respectively. μ_{ν} was the most influential parameter that determined the model output, followed by parameters Nv, k_2 , k_1 , and r.

parameter sets were tested to simulate the model output V. using the time integrated AUC value. The data files "HIVexample.txt" (lower and upper bound of parameters), "HIVparameters.txt" (generated parameter "HIVresults.txt" (model output simulation files), and "HIVSobolIndices.txt" (sensitivity analysis results) can be found in the attached Supplementary Materials S1-S13, which are available online. As shown in Figure 6, μ_{ν} was the most influential parameter followed by N_V and k_2 ; parameters k_1 and r are marginally significant for determining the model output. The model of time duration (4,000 d) was also simulated, with the same conclusion reached. Our results indicate that the second-order interaction of these parameters as $\mu_T^*\mu_N$ $k_2^*\mu_N$ and $Nv^*\mu_V$ also contribute to the variance of model output V.

These results are consistent with those from the previous publication that used eFAST technique, indicating that μ_{ν} , $N_{V_{\nu}}$ and k_{2} were the key parameters that determined the free virus V output. The author also conducted a partial rank correlation coefficient sensitivity analysis to determine the important parameters, and claimed that the overall results are not consistent between partial rank correlation coefficient and eFAST. Our results, on the other hand, indicate that there is agreement between eFAST and Sobol sensitivity analysis methods.

Enhanced PK/PD model: identification of the drug target

In our final example, we will use our previously published enhanced PK/PD model that links sunitinib plasma pharmacokinetics (PK, c.f. case example 1) with the systems biology model (PD) of vascular endothelial growth factor receptor (VEGFR) mediated cell proliferation, survival and cell permeability signaling.¹¹ Sobol sensitivity analysis was applied in our study to identify the key targets that will enhance anti-cancer efficacy when given in combination with sunitinib.

The VEGF-VEGFR mediated signaling transduction and propagation was modeled using ordinary differential equations with the cellular outputs of ERK and Akt phosphorylation and activation (pERK and pAkt). The biological

mechanism of reaction species' synthesis and degradation, as well as feedback inhibitions were all incorporated into the complex model. The steady state of both pERK and pAkt showed sustained response patterns to the ligand VEGF stimulation: sunitinib plasma concentration drove the VEGFR mediated network responses through competitive inhibition of the receptor VEGFR, where both pERK and pAkt kinetics showed transient inhibition patterns. 11 In order to understand important steps in determining the sunitinib inhibition of both pERK and pAkt, and to identify the "druggable" target, Sobol sensitivity analysis was performed. The data files "Parameterslu.txt" (lower and upper bound of parameters), "ParaSobolsets.txt" (generated parameter sets) (this file size is too big for attachment and it can be provided upon request), "AktPoutput.txt" (model output simulation files), and "AktPSobolIndices.txt" (sensitivity analysis results) can be found in the online supplemental material.

Similar to other target identification studies, the area between baseline and effect curve was used as the model output for determining the changes in pAkt over time. 11,21,51 As shown in our previous work, there were seven important parameters and steps associated with pAkt inhibition by sunitinib. The PI3K-catalyzed PIP2 to PIP3 reaction (K_{m9} , k_{g} , k_{g} , and k_{g}) supported a PI3K inhibitor in enhancing sunitinib inhibition. 11 As results of the Sobol sensitivity analysis, a PI3K inhibitor can be used to improve the efficacy in together with sunitinib usage.

In order to test the sensitivity analysis results and the model predictive performances, both pERK and pAkt inhibition was predicted when given Sunitinib alone, a PI3K inhibitor alone, and sunitinib & a PI3K inhibitor in combination. When sunitinib and a PI3K inhibitor was used in combination, pERK was inhibited to around \sim 40% at the nadirs, and the duration of complete inhibition of pAkt increased; in contrast, Sunitinib alone caused around \sim 45% inhibition at the respective nadirs of both pERK and pAkt; a PI3K inhibitor alone caused around \sim 65% and \sim 100% inhibition of both pERK and pAkt at their respective nadirs (**Figure 7**).

There are also other examples of systems pharmacology models in the literature that make use of sensitivity analysis in order to determine which parameter or which combination of parameters is determining the model output of interest. For example, for the signal transduction and activation of transcription-3 (Stat3), two important parameter interactions, i.e., the nuclear export and docking of Stat-3 to the activated receptor, were identified using sensitivity analysis.⁵² Although most of the sensitivity analysis focused on the intracellular signaling pathway at a single scale, more and more multiscale systems pharmacology models are developed to understand the drug action at both the cellular and multi-cellular levels. Throughout the tutorial we focused on the influence of model parameters and their interactions on model output. However, for future analyses it will be also useful to consider the impact of initial model conditions on its output given that the concentrations of reaction species can vary across different cells and tissues within physiological limits. 52-57

In summary, Sobol sensitivity analysis is a useful tool for model building and evaluation and thus for translational

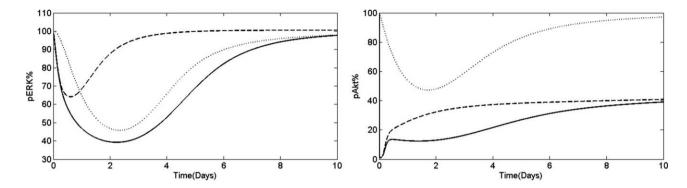


Figure 7 The model output pERK (left) and pAkt (right) inhibition in presence of drug sunitinib (gray dotted lines), a PI3K inhibitor (black dashed lines), and sunitinib and a PI3K inhibitor used in combination (solid black lines).

drug research and development. Sobol sensitivity analysis can be applied to PK, physiologically based pharmacokinetic (PBPK), and systems pharmacology models to identify the key parameters that drive the model output. In addition, Sobol sensitivity analysis can be used in a systems pharmacology context to guide target identification, which becomes particularly important for novel drug discovery and for identifying optimal drug combinations.

We hope that the conceptual discussion of this innovative analysis technique in combination with the case examples will provide the reader with a useful guide for applying Sobol sensitivity analysis to their PK, PK/PD, ePK/PD, PBPK and system pharmacology problems.

Acknowledgments. We thank Jonathan Herman at Cornell University for all the useful discussions and software for Sobol sensitivity analysis.

Conflict of Interest. The authors declared no conflict of interest.

- Rogers, M., Lyster, P. & Okita, R. NIH support for the emergence of quantitative and systems pharmacology. CPT Pharmacometrics Syst. Pharmacol. 2, e37 (2013).
- del Sol, A., Balling, R., Hood, L. & Galas, D. Diseases as network perturbations. Curr. Opin. Biotechnol. 21, 566–571 (2010).
- 3. Kitano, H. Computational systems biology. Nature. 420, 206-210 (2002).
- Schmidt, S., Post, T.M., Peletier, L.A., Boroujerdi, M.A. & Danhof, M. Coping with time scales in disease systems analysis: application to bone remodeling. *J. Pharma-cokinet. Pharmacodyn.* 38, 873–900 (2011).
- Lesko, L.J., Zheng, S. & Schmidt, S. Systems approaches in risk assessment. Clin. Pharmacol. Ther. 93, 413–424 (2013).
- van Riel, N.A.W. Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief. Bioinform.* 7, 364–374 (2006).
- Breimer, D.D. & Danhof, M. Relevance of the application of pharmacokineticpharmacodynamic modelling concepts in drug development. The "wooden shoe" paradigm. Clin. Pharmacokinet. 32, 259–267 (1997).
- Danhof, M., de Jongh, J., De Lange, E.C., Della Pasqua, O., Ploeger, B.A. & Voskuyl, R.A. Mechanism-based pharmacokinetic-pharmacodynamic modeling: biophase distribution, receptor theory, and dynamical systems analysis. *Annu. Rev. Pharmacol. Toxicol.* 47, 357–400 (2007).
- Birtwistle, M.R., Hatakeyama, M., Yumoto, N., Ogunnaike, B.A., Hoek, J.B. & Kholodenko, B.N. Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol. Syst. Biol.* 3, 144 (2007).
- Iyengar, R., Zhao, S., Chung, S.W., Mager, D.E. & Gallo, J.M. Merging systems biology with pharmacodynamics. Sci. Transl. Med. 4, 126–127 (2012).
- Zhang, X.Y., Birtwistle, M.R. & Gallo, J.M. A general network pharmacodynamic model-based design pipeline for customized cancer therapy applied to the VEGFR pathway. CPT Pharmacometrics Syst. Pharmacol. 3, e92 (2014).
- Jusko, W.J. Moving from basic toward systems pharmacodynamic models. J. Pharm. Sci. 102, 2930–2940 (2013).

- Benson, N., Metelkin, E., Demin, O., Li, G.L., Nichols, D. & van der Graaf, P.H. A systems pharmacology perspective on the clinical development of fatty acid amide hydrolase inhibitors for pain. CPT Pharmacometrics Syst. Pharmacol. 3, e91 (2014).
- Brannmark, C., Nyman, E., Fagerholm, S., Bergenholm, L., Ekstrand, E.M., Cedersund, G. & Stralfors, P. 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).
- Nestorov, I., Graham, G., Duffull, S., Aarons, L., Fuseau, E. & Coates, P. Modeling and stimulation for clinical trial design involving a categorical response: a phase II case study with naratriptan. *Pharm. Res.* 18, 1210–1219 (2001).
- Freeman, M. Feedback control of intercellular signalling in development. Nature 408, 313–319 (2000).
- Bhalla, U.S. & Iyengar, R. Emergent properties of networks of biological signaling pathways. Science 283, 381–387 (1999).
- Agoram, B. Evaluating systems pharmacology models is different from evaluating standard pharmacokinetic-pharmacodynamic models. CPT Pharmacometrics Syst. Pharmacol. 3, e101 (2014).
- Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R. & Sethna, J.P. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3, 1871–1878 (2007).
- Vanlier, J., Tiemann, C.A., Hilbers, P.A. & van Riel, N.A. Parameter uncertainty in biochemical models described by ordinary differential equations. *Math. Biosci.* 246, 305–314 (2013).
- Benson, N. et al. Systems pharmacology of the nerve growth factor pathway: use of a systems biology model for the identification of key drug targets using sensitivity analysis and the integration of physiology and pharmacology. Interface Focus 3, 20120071 (2013).
- Sarkar, A.X. & Sobie, E.A. Quantification of repolarization reserve to understand interpatient variability in the response to proarrhythmic drugs: a computational analysis. Heart Rhythm 8, 1749–1755 (2011).
- Cummins, M.A., Dalal, P.J., Bugana, M., Severi, S. & Sobie, E.A. Comprehensive analyses of ventricular myocyte models identify targets exhibiting favorable rate dependence. *PLoS Comput. Biol.* 10, e1003543 (2014).
- Saltelli, A., Chan, K. & Scott, E.M. Sensitivity Analysis: Wiley Series in Probability and Statistics (John Wiley, Chichester, 2000).
- Kent, E., Neumann, S., Kummer, U. & Mendes, P. What can we learn from global sensitivity analysis of biochemical systems? PLoS One 8, e79244 (2013).
- Sahle, S., Mendes, P., Hoops, S. & Kummer, U. A new strategy for assessing sensitivities in biochemical models. *Philos. Trans. R. Soc. A.* 366, 3619–3631 (2008).
- Frey, H.C. & Patil, S.R. Identification and review of sensitivity analysis methods. *Risk Anal.* 22, 553–578 (2002).
- 28. Saltelli, A. Global Sensitivity Analysis: The Primer (John Wiley, Chichester, 2008).
- Mokhtari, A. & Frey, H.C. Sensitivity analysis of a two-dimensional probabilistic risk assessment model using analysis of variance. Risk Anal. 25, 1511–1529 (2005).
- Kiparissides, A., Kucherenko, S.S., Mantalaris, A. & Pistikopoulos, E.N. Global sensitivity analysis challenges in biological systems modeling. *Ind. Eng. Chem. Res.* 48, 7168–7180 (2009).
- Hamby, D.M. A review of techniques for parameter sensitivity analysis of environmental models. Environ. Monit. Assess. 32, 135–154 (1994).
- Ingalls, B. Sensitivity analysis: from model parameters to system behaviour. Essays Biochem. 45, 177–193 (2008).
- Zi, Z. Sensitivity analysis approaches applied to systems biology models. IET Syst. Biol. 5, 336–336 (2011).
- Zi, Z., Zheng, Y., Rundell, A.E. & Klipp, E. SBML-SAT: a systems biology markup language (SBML) based sensitivity analysis tool. BMC Bioinformatics 9, 342 (2008).
- Marino, S., Hogue, I.B., Ray, C.J. & Kirschner, D.E. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Biol. 254, 178–196 (2008).

- Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* 55, 271–280 (2001).
- Bentele, M. et al. Mathematical modeling reveals threshold mechanism in CD95induced apoptosis. J. Cell Biol. 166, 839–851 (2004).
- Blower, S.M. & Dowlatabadi, H., Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int. Stat. Rev.* 62, 229–243 (1994).
- Zi, Z., Cho, K.H., Sung, M.H., Xia, X., Zheng, J. & Sun, Z. In silico identification of the key components and steps in IFN-gamma induced JAK-STAT signaling pathway. FEBS Lett. 579, 1101–1108 (2005).
- Cho, K.H., Shin, S.Y., Lee, H.W. & Wolkenhauer, O. Investigations into the analysis and modeling of the TNF alpha-mediated NF-kappa B-signaling pathway. *Genome Res.* 13, 2413–2422 (2003).
- 41. Yoon, J. & Deisboeck, T.S. Investigating differential dynamics of the MAPK signaling cascade using a multi-parametric global sensitivity analysis. *PLoS One* **4**, e4560 (2009).
- Saltelli, A., Tarantola, S., & Chan, KPS. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*. 41, 39–56 (1999).
- Mokhtari, A., Christopher Frey, H. & Zheng, J. Evaluation and recommendation of sensitivity analysis methods for application to Stochastic Human Exposure and Dose Simulation models. J. Expo. Sci. Environ. Epidemiol. 16, 491–506 (2006).
- Sobol, I.M. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comp. Exp. 1, 407–414 (1993).
- Chan, K., Saltelli, A. & Tarantola, S. Sensitivity analysis of model output: variancebased methods make the difference. Proceedings of the 1997 Winter Simulation Conference, Atlanta, Georgia, USA (1997).
- Sobol, I.M. On the distribution of points in a cube and the approximate evaluation of integrals. USSR Comp. Math. Math+. 7, 86–112 (1967).
- Dalal, I.L., Stefan, D. & Harwayne-Gidansky, J. Low discrepancy sequences for Monte Carlo simulations on reconfigurable platforms. International Conference on Application-Specific Systems, Architectures and Processors, Proceedings, Leuven, Belgium (2008).
- Houk, B.E., Bello, C.L., Kang, D. & Amantea, M. A population pharmacokinetic meta-analysis of sunitinib malate (SU11248) and its primary metabolite (SU12662) in healthy volunteers and oncology patients. *Clin. Cancer Res.* 15, 2497–2506 (2009).

- Gueorguieva, I., Nestorov, I.A., Aarons, L. & Rowland, M. Uncertainty analysis in pharmacokinetics and pharmacodynamics: application to naratriptan. *Pharm. Res.* 22, 1614–1626 (2005).
- Perelson, A.S., Kirschner, D.E. & De Boer, R. Dynamics of HIV infection of CD4+ T cells. Math. Biosci. 114, 81–125 (1993).
- Lebedeva, G. et al. Model-based global sensitivity analysis as applied to identification of anti-cancer drug targets and biomarkers of drug resistance in the ErbB2/3 network. Eur. J. Pharm. Sci. 46, 244–258 (2012).
- Mahdavi, A., Davey, R.E., Bhola, P., Yin, T. & Zandstra, P.W. Sensitivity analysis of intracellular signaling pathway kinetics predicts targets for stem cell fate control. *PLoS Comput. Biol.* 3, e130 (2007).
- 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).
- Peterson, M.C. & Riggs, M.M. A physiologically based mathematical model of integrated calcium homeostasis and bone remodeling. Bone 46, 49–63 (2010).
- de Pillis, L.G., Radunskaya, A.E. & Wiseman, C.L. A validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res.* 65, 7950–7958 (2005).
- Wilkinson, D.J. Stochastic modelling for quantitative description of heterogeneous biological systems. Nat. Rev. Genet. 10, 122–133 (2009).
- Gowal, S., de Giacomi, M. & Le Boudec, J.Y. Comment on: a validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res.* 67, 8419–8421; author reply 8420-8411 (2007).

© 2015 The Authors CPT: Pharmacometrics & Systems Pharmacology published by Wiley Periodicals, Inc. on behalf of American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.wileyonlinelibrary.com/psp4)