MASTER THESIS COMPUTER SCIENCE

Exploring Practical Identifiability in Target-Mediated Drug Disposition Models

Investigation of the influence of experimental conditions on identifiability of a series of target-mediated drug disposition models through simulation

Author:

Helle VAN DEN MAAGDENBERG,
s1828231

Supervisors:

Lu CAO, LIACS
Tamara VAN STEEG, LAP&P
Richard HOOIJMAIJERS, LAP&P

June 2021

Universiteit Leiden
The Netherlands

LAP Leiden
Advanced PK-PD

Contents

1 Abstract

Accurately determining parameter identifiability is a continuing issue within the field of pharmacometrics. More specifically, it is a challenge for drugs with high target affinity and specificity displaying target-mediated drug disposition (TMDD) kinetics. TMDD models are often over-parameterized and difficult to converge.
In this study, a workflow for determining parameter identifiability in TMDD models was proposed. The workflow combines three methods: variance-based sensitivity analysis using eFAST, analysis of the eigenvalues of the Fisher-Information Matrix and log-likelihood profiling. The workflow was partly evaluated by testing the Fisher-Information Matrix and the variance-based sensitivity approach on simulated data sets. The simulated data sets were also used to learn about the identifiability of parameters in TMDD models under different conditions.
The simulated data sets were varied in parameter values and biomarkers. A full two-compartment TMDD model and two simplifications (Quasi-steady state approximation and Michaelis-Menten approximation) were fitted on these data sets with NONMEM. Regarding identifiability under different conditions, the most remarkable result was that fitting a data set with free receptor concentration, a Michaelis-Menten approximation and a low value of target-complex degeneration (k_{int}), resulted in strong overestimation of the k_{int}. This is caused by the disagreement between the true TMDD profile and the Michaelis-Menten profile for low values of k_{int}. The parameter estimates from NONMEM models fitted on the simulated data set were compared to predictions of the Fisher-Information Matrix and the variance-based sensitivity approach (eFAST). The results from this analysis showed that the precision for both the Fisher-Information Matrix (20%) and the variance-based sensitivity (23%) method were low, indicating that neither of these approaches are reliable for determining parameter identifiability in TMDD models. However, further research and possible adjustments are needed to make any definitive conclusions on the usefulness of this workflow for determining parameter identifiability in TMDD models.

2 Acknowledgments

Firstly, I would like to sincerely thank my three supervisors: Tamara van Steeg for all her good advice and assistance every step of the way; Richard Hooijmaijers for his help with my code and every other problem I ran into; Lu Cao for her continuous support throughout this project. Likewise, I express my gratitude to Martijn van Noort for sharing his ideas and advice on the project and his seemingly endless patience in explaining the relevant mathematical concepts. I would also like to acknowledge Martijn Ruppert for his advice on the statistical analysis of the results. In addition, thank you to David Fairman for his constructive feedback on the project and assistance in selecting realistic parameter values. Finally, I extend my thanks to Henk-Jan Drenth and the LAP&P for providing me with the fantastic opportunity to work on this project and everyone there whom I haven't mentioned specifically for their interest, questions, advice and support.

3 Introduction

3.1 Pharmacological modeling

It is imperative to study the relationship between drug concentration, response and its underlying processes to understand and predict drug effects. This relationship is dependent on many factors such as absorption, distribution, metabolism and excretion of the drug, but also the interaction between drug and target. The study of the fate of a drug in the body is called pharmacokinetics (PK) and the study of how the body reacts to a drug is called pharmacodynamics (PD). To predict effects and concentration in new situations, such as different individuals, doses or administration, pharmacokinetic/pharmacodynamic (PK/PD) models are used. A method that is often employed for constructing PK/PD models is nonlinear mixed-effects modeling [1]. These models consist of fixed and random effects. The fixed effects capture the typical changes over time in the measured compounds and/or biomarkers. The random effects describe the variability within the study population, which includes two different types of variability. The variability between individuals is called the inter-individual variability and the other is the residual unexplained variability. A software program widely used for developing PK/PD models is NONMEM (NON-linear Mixed Effects Modeling) [2].

3.2 Target Mediated Drug Disposition

Some drugs exhibit non-linear kinetics. A possible reason for non-linear kinetics is that the drug has a high target affinity in combination with a high target specificity and limited target capacity. The kinetics and dynamics of these drugs can be described by Target-Mediated Drug Disposition (TMDD) models. Target-Mediated Drug Disposition was first described by Levy [3]. In this article the principles of TMDD were illustrated using anti-coagulant Warfarin, which is a small molecule drug. However, with the rapid development of protein drugs the interest in TMDD grew. Protein drugs are large molecule drugs which often display high target affinity and specificity, making them ideal drug candidates. Due to their high target affinity, they regularly demonstrate the non-linear kinetics that are described by TMDD models. [4]
The full TMDD model is similar to a two-compartment model, with an extension consisting of compartments for the free drug, the target and the associated drug-target complex with rate constants for the association and dissociation of the drug and its target. So this model actually includes both the pharmacokinetics and the pharmacodynamics. Ideally, biomarkers in all the compartments are measured, however often only the drug compound is measured. A schematic representation can be found in figure 2. The TMDD kinetics were first written as set of ordinary differential equations by Mager and Jusko [5] for a one-compartment model and later expanded to a two-compartment model by Gibiansky [6].

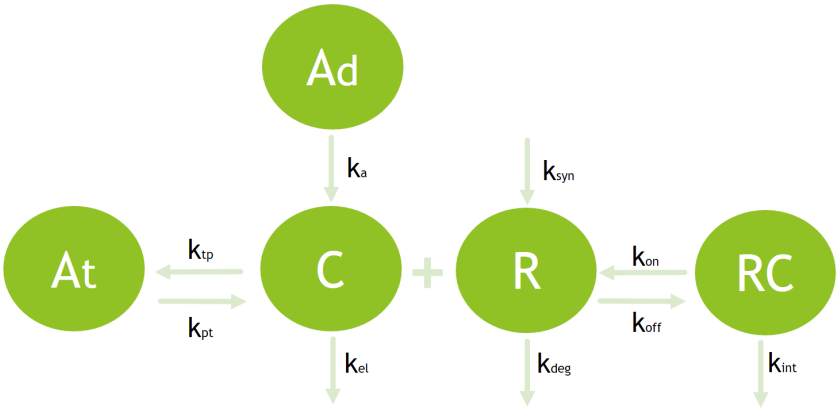

Figure 2: Schematic representation of a target-mediated drug disposition PK/PD model. Circles represent the compartments and arrows are the rate constants. Detailed description and equations can be found in the methods.

A typical pharmacokinetic profile of TMDD can be seen in figure 3. This concentration-time profile contains four phases [7].

1. (A) The fast concentration drop in the short initial phase is mainly caused by the fast association of drug and target.
2. (B) The second phase is a linear decline of the concentration caused by the clearance of free drug from the central compartment.
3. (C) The third phase is a non-linear transition phase with mixed elimination of free drug and internalization of drug-target complex.
4. (D) Eventually, most of the free drug has been eliminated. The fourth phase is another linear decline phase which is characterized by the slow decrease of drug concentration mainly through internalization of the drug-target complex.

The phases are not always as pronounced as seen in figure 3 and are strongly dependent on the dose and parameter values.

3.3 Simplifications of the full TMDD model

The full TMDD model is often over-parameterized and difficult to converge. A solution is to fix certain parameters or the model can be simplified. Numerous simplifications of the TMDD model were proposed [5, 6, 8, 9]. A few of the commonly used simplifications are explained in more detail below.

3.3.1 Quasi-steady state assumption

The first simplification that will be discussed is the quasi-steady state (QSS) assumption, which was proposed by Gibiansky [6]. In this simplification of the TMDD model equations, the assumption is made that the association (k_{on}) and disassociation (k_{off}) & complex degradation/internalization (k_{int}) are in a quasi-steady state. This assumption is made as in general these processes are much faster compared to the others. The quasi-steady state constant Kss, which

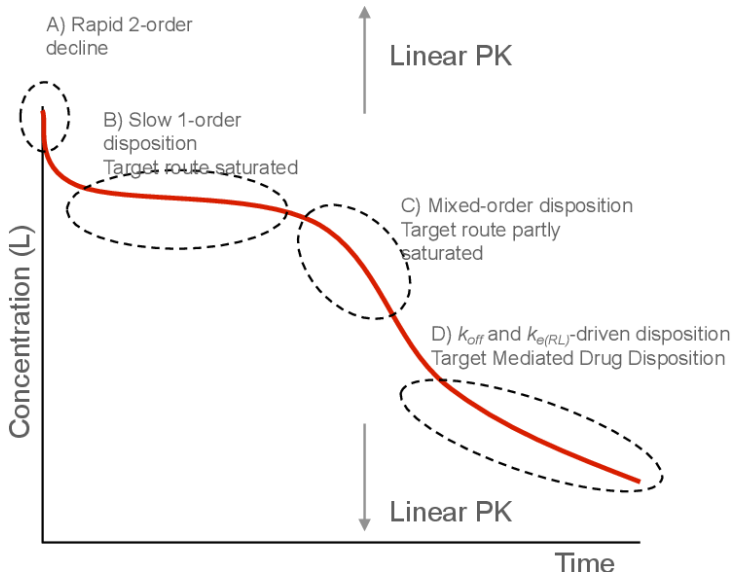

Figure 3: Graph showing the typical four phases of the concentration-time profile of a drug displaying target-mediated drug disposition kinetics. Image from Peletier and Gabrielsson (2012) [7]

represents this balance, is added to the ordinary differential equations [6]. As k_{on} and k_{off} are not used in the equations of the QSS, this simplification reduces the number of parameters by one.

3.3.2 Quasi-Equilibrium

Another common simplification similar to the QSS assumption is the Quasi-Equilibrium (QE) assumption [8]. In this case the equations are the same as for the QSS assumption, but with the K_{ss} replaced by K_d, which is calculated as k_{off}/k_{on}. This simplification is valid if the rate of internalization of the drug-target complex (k_{int}) is negligible compared to the dissociation rate. In practice, these two simplifications are the same when fitting these models, only the underlying parameterization is different.

3.3.3 Michaelis-Menten approximation

The Michaelis-Menten (MM) approximation describes the drug binding kinetics of the TMDD model in Michealis-Menten terms. The constant V_{max} and the equilibrium constant K_m are used. K_m is equal to K_{ss} from the QSS assumption. V_{max} is equal to total receptor concentration (R_{tot}) * k_{int}. The additional assumption is made here that R_{tot} is constant, therefore V_{max} is a constant as well and the derivative of R_{tot} is 0 [6].

3.4 Parameter Identifiability

Selecting the right model for predicting (adverse) effects and determining appropriate dosing of drugs is a difficult task. An oversimplified model may limit usability or even result in inaccurate predictions. However, a too complex model may suffer from over-fitting issues and may result in parameter values that do not extrapolate well to new data. If parameter values cannot be uniquely determined this is called parameter

unidentifiability. Therefore, analyzing identifiability is the subject of many scientific publications, both within the scope of pharmacology as well as in other research domains, see e.g. [6, 10–18].

3.5 Parameter Identifiability Analysis

As a full target-mediated drug disposition model is often over-parameterized, it is important to determine when and how to simplify. Choosing which parameters to fix or which simplification to use can be determined by parameter identifiability analysis. Listed as follows are the main concepts pertaining to parameter identifiability.

- **Structural Identifiability**; parameter values can be determined given an infinite amount of noise-free data.
- **Practical Identifiability**; parameter values can be accurately estimated given a finite amount of noisy experimental data.
- **Global Identifiability analysis**; analysis of the identifiability of a parameter over the entire possible parameter range.
- **Local identifiability analysis**; analysis of the identifiability of a parameter over a limited parameter range.
- **Globally identifiable**; the parameter is identifiable over the entire possible parameter range.
- **locally identifiable**; the parameter is identifiable at any point in the possible parameter range.

The definitions for identifiability are based on those given by Browning et al. [19]. Structural parameter identifiability is a prerequisite for successful parameter estimation. However, as in practice there is never an infinite amount of noise free data, the main interest in this study is practical identifiability. Especially, in phase 1 studies which generally use a small number of study subjects, there is a limited data set. Phase 1 is the first test of a new drug candidate in humans and is used to assess safety in healthy individuals and to determine the PK and PD. Several approaches have been proposed in literature for analyzing practical parameter identifiability, a few of which are described below.

3.5.1 Gibiansky Hierarchical approach

In Gibiansky [6] an algorithm for determining parameter identifiability is proposed. First the full TMDD model is fitted to the data and subsequently the parameter estimates are used for fitting increasingly simplified models in this order QSS, QE, MM approximations. These four models are then used for simulating concentration-time profiles, which are compared to the full model. The simplest model that is equivalent to the full model within the range of the available data is selected. The reason for this is that if a more complex models is equivalent to simpler models, the extra added parameters in the more complex model are not reliable. as they have no influence on the result and are thus not identifiable. As fitting the more complex

models, especially the full TMDD model, is sometimes unstable, the unfittable model is skipped in this approach. Sometimes it helps to use the parameter estimates from a simpler model as initial estimates for the full model, to stabilize the fitting.

3.5.2 Bootstrapping

Bootstrapping is one of the most reliable methods of estimating parameter identifiability, but it is also very computationally demanding. This method consists of repeatedly resampling from the observations and refitting the model on this data. The width of the distribution of parameter estimates represents the reliability of the parameter estimates. Different methods for resampling are possible. [20] In Thai et al. [20] it was found that bootstrapping is a suitable method for determining uncertainty in non-linear mixed effects models.

3.5.3 Profile Likelihood

The profile likelihood identifiability analysis consists of fixing a parameter over a range of values and fitting all the remaining parameters, calculating the objective function value for all the fits. The objective function is a measure of how well a model fits the observed data. In other words, it is the likelihood of the parameter vector given the data. Then the likelihood profile is constructed from the objective function values for all the fitted parameter sets. If a parameter is unidentifiable, its value will have no or little effect on the objective function values, which will result in a shallow or even flat profile. This approach is feasible for both structural and practical identifiability analysis. [21]. This approach has two major disadvantages. Firstly, it is very computationally intensive, albeit not as much as the bootstrapping method. Secondly, as it is a univariate test it assumes no covariance between variables, which may not be valid in all cases. Therefore, an extension to log-likelihood profiling (LLP) is proposed in [11], where they combine it with sampling importance resampling (SIR). SIR was first proposed as a method for determining estimates of uncertainty distributions in Non-linear Mixed-effects modeling by [18]. The first step in SIR is sampling from a proposal distribution, which in case of [11] is the log-likelihood profile. Secondly, the importance ratios are calculated from the samples. Finally, through resampling using the importance ratios the confidence intervals and standard errors of the parameters can be determined.

3.5.4 Fisher Information Matrix

One of the most common ways to assess parameter identifiability is through studying the standard errors of parameter estimates. The standard errors are calculated in the covariance step in NONMEM by taking the inverse of the fisher information matrix (FIM), which is the covariance matrix [2]. However, the covariance step in NONMEM often fails, which can be caused by a non-invertible FIM, where the determinant of the matrix is zero.
Fisher Information provides knowledge about how much information data holds about the true parameter value. The FIM is a matrix of n x n, where n is the number of parameters. The values in the FIM represent the variance around the current parameter estimate. Thus the FIM and its covariance matrix can also be used directly to find identifiable parameter combinations [10]. The parameters in these

combinations are unidentifiable individually, but are identifiable together. This is valuable information for deciding on which simplification to choose in case of an unidentifiable model.
Another possible method could be to study the eigenvalues and vectors of the Fisher Information Matrix, so the covariance matrix is unnecessary. The eigenvalues indicate the size of the effect on the output of a step in the direction of the eigenvector. Therefore, the eigenvector of a large eigenvalue is an "informative direction" and the eigenvector of a small eigenvalue might indicate a unidentifiable direction. As the direction of eigenvector can be decomposed by parameter, the parameters that "contribute" the most to the direction can be marked as unidentifiable parameter (combinations).

3.5.5 Variance-based sensitivity analysis

Sensitivity analysis is a technique for determining parameter identifiability. A parameter is sensitive if a change in its value results in a change in model output. If this is the case, the parameter is identifiable. In variance-based sensitivity analysis parameter values are incrementally varied, so their influence on model output can be determined. To describe the variance in output in different dimensionality Sobol indices were introduced. Sobol indices are a quantitative measure of sensitivity [15, 16]. The first order sensitivity index reflects the direct influence of a single parameter on the output. Higher order indices reflect the influence of the interaction between sets of parameters on the output (second order is influence of interaction between two parameters, third order between three, etc.). Both the first order and interaction term together is the total order sensitivity. This is the total influence of a parameter on the output. The Sobol indices can be calculated using extended Fourier amplitude sensitivity analysis (eFAST) [14]. Performing variance-based sensitivity analysis using eFAST, is a method of global sensitivity analysis. In a comparison between different global sensitivity analysis methods for finding influential parameters in physiologically-based pharmacokinetic modeling, it was found that calculating Sobol indices uses eFAST provided the best balance between reliability and efficiency [13].

3.6 Proposed workflow for determining parameter identifiability in TMDD models

All the described methods for determining parameter identifiability have there own advantages and disadvantages. This suggests that combining these methods gives a more reliable results. In figure 4 a suggested workflow is proposed which combines three identifiability methods that guide the modelling process. Some inspiration is taken from the hierarchical method [6], by starting with the most complex model and then iteratively simplifying the model. However, because the simplification steps are guided by parameter specific methods, choices can also be made between "hierarchically equal" models, making this method more flexible and widely applicable. In the first step of the process the eFAST method [14] is utilized as a global sensitivity analysis. This method applies only to the structural part of the model. It is a first scan for unidentifiable model parameters. If such a model parameter is found, the model is simplified by either fixing this parameter or choosing a less complex model. The

sensitivity analysis is then performed again, until no more unidentifiable parameters are found.
Subsequently if no un-identifiable parameters are found, the model is fitted with NONMEM. The resulting parameter estimates can be used to apply the Fisher Information Matrix approach. This is a local method of identifiability, however it also applies to the random effects in addition to the fixed effects. Another advantage is that it could, in theory, be used to find identifiable parameter combinations, further aiding in the selection of simplifications [10]. If non-identifiable parameter (combinations) are found, the process will be initiated again from the global sensitivity analysis. If not, a final check of sensitivity is done using the profile likelihood on parameters which are often misclassified as identifiable by the other two methods. This solves one of the disadvantages of profile likelihood, namely the computational intensity of the method, as it now only has to be applied to a few parameters. This is the only one of these methods which uses the actual experiment data directly.

3.7 Study goals

This study has three main goals. The first is to gather more information about the identifiability of TMDD model parameters with different parameter values through fitting different models with NONMEM on simulated data sets consisting of different biomarkers. Secondly to test the reliability of a part of the proposed workflow, namely the eFAST test and the Fisher Information Matrix test. Finally, identify the "known to be often misclassified" parameters for confirmation by the profile likelihood part of the workflow.

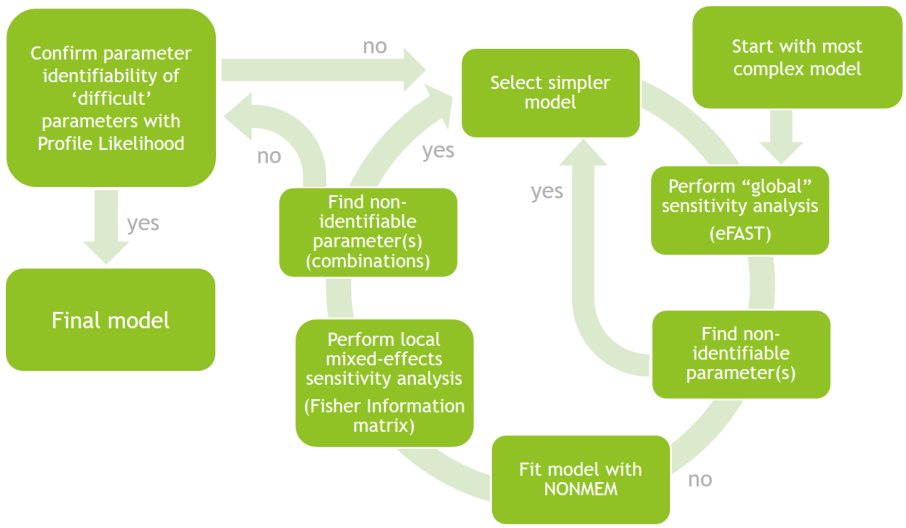

Figure 4: Proposed workflow for determining parameter identifiability in TMDD models.

4 Methods

4.1 Software

All simulations and analyses were performed in Rstudio 1.3.1093 [22] using R 4.0.3 [23]. Models were fitted in NONMEM 7.5.0 [2] using perl-speaks-NONMEM (psN) [24] in parallel controlled from R.

4.2 Ordinary differential equations models

The ordinary differential equations and parameter derivations (equations 1-12) of a two-compartment TMDD model with oral dosing adapted from [6] are given below.

$$k_{el} = \frac{Cl}{V1} \tag{1}$$

$$k_{pt} = \frac{Q}{V1} \tag{2}$$

$$k_{tp} = \frac{Q}{V2} \tag{3}$$

$$k_{syn} = k_{deg}R0 \tag{4}$$

$$\frac{d(A_d)}{dt} = -k_a A_d \tag{5}$$

$$\frac{d(C)}{dt} = \frac{k_a A_d}{V1} - (k_{el} + k_{pt})C - k_{on}C \cdot R + k_{off}RC + k_{tp}\frac{A_t}{V1} \tag{6}$$

$$\frac{d(A_t)}{dt} = k_{pt}C \cdot V1 - k_{tp}A_t \tag{7}$$

$$\frac{d(R)}{dt} = k_{syn} - k_{deg}R - k_{on}C \cdot R + k_{off}RC \tag{8}$$

$$\frac{d(RC)}{dt} = k_{on}C \cdot R - (k_{int} + k_{off})RC \tag{9}$$

$$C_{tot} = C + RC \tag{10}$$

$$R_{tot} = R + RC \tag{11}$$

$$A_d(0) = D1; C(0) = \frac{D2}{V1}; A_t(0) = 0; RC(0) = 0; R(0) = \frac{k_{syn}}{k_{deg}} \tag{12}$$

C is the free drug concentration, R refers to the free target concentration and RC is the concentration of the drug-target complex (the product). Cl is the clearance. Q is the inter-compartmental clearance. $V1$ is the volume of distribution of the central compartment and $V2$ of the peripheral compartment. k_{el} is the rate of elimination of the drug. k_{syn} is the rate of influx/synthesis of the target and k_{deg} is the rate of degradation of the target. k_{on} is the rate of association and k_{off} is the rate of dissociation of the drug-target complex. k_{tp} and k_{pt} are diffusion to and from the peripheral compartment. k_{int} is the rate of internalization of the target complex. Equation 12 gives the initial states of the compartments, with $D1$ a subcutaneous dose and $D2$ an intravenous dose.
The quasi-steady state constant (K_{ss}) is used to represent the balance between k_{int}, k_{on} and k_{off} in the Quasi-Steady state approximation. It is calculated as follows:

$$K_{ss} = \frac{k_{int} + k_{off}}{k_{on}} \tag{13}$$

The ordinary differential equations and derived parameters are as follows:

$$k_{el} = \frac{Cl}{V1} \tag{14}$$
$$k_{pt} = \frac{Q}{V1} \tag{15}$$
$$k_{tp} = \frac{Q}{V2} \tag{16}$$
$$k_{syn} = k_{deg} R0 \tag{17}$$
$$\frac{d(A_d)}{dt} = -k_a A_d \tag{18}$$
$$\frac{d(C_{tot})}{dt} = \frac{k_a A_d}{V1} - (k_{el} + k_{pt})C - \frac{R_{tot} k_{int} C}{K_{ss} + C} + \frac{k_{tp} A_t}{V1} \tag{19}$$
$$\frac{d(A_t)}{dt} = k_{pt} C \cdot V1 - k_{tp} A_t \tag{20}$$
$$\frac{d(R_{tot})}{dt} = k_{syn} - k_{deg} R_{tot} - (k_{int} - k_{deg}) \frac{R_{tot} C}{K_{ss} + C} \tag{21}$$
$$C = \frac{1}{2}\left[(C_{tot} - R_{tot} - K_{ss}) + \sqrt{(C_{tot} - R_{tot} - K_{ssS})^2 + 4K_{ss} C_{tot}}\right] \tag{22}$$
$$RC = \frac{R0 \cdot C}{Kss + C} \tag{23}$$
$$R = R_{tot} - RC \tag{24}$$
$$A_d(0) = D1; C_{tot}(0) = \frac{D2}{V1}; A_t(0) = 0; R_{tot}(0) = \frac{k_{syn}}{k_{deg}} \tag{25}$$

The Michaelis-Menten approximation describes the drug binding kinetics of the TMDD model in Michealis-Menten terms. [6]. K_m is equal to the K_{ss} in the QSS equations and represents the Michaelis-Menten constant of the non-linear elimination. V_{max} is the maximum non-linear elimination.

$$V_{max} = R_{tot} k_{int} \tag{26}$$
$$k_m = \frac{k_{int} + k_{off}}{k_{on}} \tag{27}$$

The ordinary differential equations and derived parameters are as follows:

$$k_{el} = \frac{Cl}{V1} \tag{28}$$

$$k_{pt} = \frac{Q}{V1} \tag{29}$$

$$k_{tp} = \frac{Q}{V2} \tag{30}$$

$$\frac{d(A_d)}{dt} = -k_a A_d \tag{31}$$

$$\frac{d(C)}{dt} = \frac{k_a A_d}{V1} - (k_{el} + k_{pt})C - \frac{V_{max} C}{K_m + C} + \frac{k_{tp} A_t}{V1} \tag{32}$$

$$\frac{d(A_t)}{dt} = k_{pt} C \cdot V1 - k_{tp} A_t \tag{33}$$

$$RC = \frac{R0 \cdot C}{Km + C} \tag{34}$$

$$R = R0 - RC \tag{35}$$

$$R_{tot} = R0 \tag{36}$$

$$C_{tot} = C + RC \tag{37}$$

$$A_d(0) = D1; C_{tot}(0) = \frac{D2}{V1}; A_t(0) = 0; \tag{38}$$

4.3 Simulated Data Sets

The simulated data sets were created using the package PopED version 0.5.0 [25, 26]. PopED is a package built for population and individual optimal experimental design. It allows for specifying experimental conditions, such as a sampling scheme, dosing scheme, number of individuals, et cetera. Moreover, the model can be specified using ordinary differential equations as well as the parameter values and the intra-individual variability and residual unexplained variability. PopED has functionality to automatically create a simulated data set from these details. The data sets created using PopED can then be easily transformed in R to a format suitable for NONMEM. For all the created data sets a full two-compartment TMDD model was used. Three different parameter value sets were used. The fixed model parameters are specified in table 1. All parameter values expect k_{deg} and k_{int} are the same between the sets. The high » 1 and low « 1 k_{int} value are chosen to simulate the different profiles of a membrane-bound and soluble receptor respectively. The higher $k_{deg} > k_{int}$ and lower $k_{deg} < k_{int}$ value were chosen, because this difference generates very different profiles.

Table 1: Parameter value sets used for generating simulated data sets

Parameter	abbreviation	units	Set 1	Set 2	Set 3
Absorption rate *	k_a	/day	0.25	0.25	0.25
Clearance	Cl	L/day	0.2	0.2	0.2
Volume of distribution central compartment	$V1$	L	3.1	3.1	3.1
Inter-compartmental clearance	Q	L/day	0.79	0.79	0.79
Volume of distribution peripheral compartment	$V2$	L	2.8	2.8	2.8
Initial receptor concentration	$R0$	nM	0.001	0.001	0.001
Rate of drug and target association	k_{on}	/nM/day	86.4	86.4	86.4
Rate of drug-target complex dissociation	k_{off}	/day	32.6	32.6	32.6
Rate of target degradation	k_{deg}	/day	**82**	**82**	**20**
Rate of drug-target complex degradation/internalization	k_{int}	/day	**0.01**	**40**	**40**

The random effects model parameters are the same over all experiments. For the inter-individual variability (η) the variance ω per parameter is given in table 3. The parameters with IIV are modeled as: $\theta * exp(\eta)$.

Table 2: Variance of the inter-individual variability (η) parameter values used for generating simulated data sets

Parameter	k_a	Cl	$V1$	$R0$	k_{deg}	$Vmax$
Omega	0.45	0.30	0.26	0.40	0.30	0.30

In table 2 the parameter values for the variance (σ) of the proportional residual unexplained variability (ϵ) is given for each biomarker.

Table 3: Residual unexplained variability parameter values used for generating simulated data sets

Parameter	C	R	RC	C_{tot}	R_{tot}
Sigma	0.250	0.300	0.400	0.325	0.350

The study design was based on a typical mAb phase 1 design in which a total of 24 individuals were included (n=6 per dose group). The simulated dose groups are 112.0 nmol, 373.6 nmol,746.4 nmol and 1493.6 nmol. The doses are administered as intravenous doses. These doses are based on the sampling scheme in [27] (0.3 mg/kg SC, 1 mg/kg SC, 2 mg/kg SC and 4 mg/kg SC), calculated from mg/kg to nmol with a molecular weight of 150 kDa (based on the MW of anrukinzumab 145.4 kDa [28]), a weight of 70kg and a bio-availability of 0.8. The sampling time points were selected as 0.5 hours, every 12 hours until day 7, once a week for 4 weeks, once every two weeks for 4 weeks and once every 4 weeks for 8 weeks. This is based on the sampling scheme in [27], considering a phase-1 study scheme design. All data sets were generated with this dosing and sampling scheme.
The sampled biomarkers were varied between the data sets. Five different types of biomarkers are used: free drug (C), free and bound drug (C_{tot}), free receptor (R), free and bound receptor (R_{tot}) and drug-receptor complex (RC). The following combinations of these biomarkers were used in the simulations:

- C
- C & R
- C, R & RC
- C & R_{tot}
- C_{tot}
- C_{tot} & R
- C_{tot} & R_{tot}

This results in 3 x 7 = 21 different simulated data sets.

4.4 NONMEM

The simulated data sets are used for fitting three different models, namely a full two-compartment TMDD model, a two-compartment QSS model and a two-compartment MM model. This gave 21 x 3 = 63 models that should be fitted with NONMEM. However, the initial parameter values in NONMEM can have a profound influence on the final results. Therefore, the choice was made to fit each of these 63 models 100 times with different initial parameter values. These 100 different initial parameter settings were generated in R by multiplying the true/simulation parameter value with a value randomly drawn from a log-normal distribution. The log-normal distribution is derived from a normal distribution with a mean of 1 and a standard deviation of 1.

4.5 Variance-based sensitivity analysis with pksensi

For each of the different structural models a variance-based sensitivity analysis was performed using the package pksensi [29] to execute the eFAST algorithm. With pksensi the eFAST algorithm can be run on every sampling point. The first order, interaction and total order measures are then the maximum value of all time points. This was extended to the different dosing groups by performing the eFAST with pksensi four times, once for each dosing group and then selecting the maximum from each time-point for each dose group. All parameters were varied between half and twice the parameter value used in the simulations. Number of samples (n) was set to 5000 and the number of replications for calculating the convergence was set to 5. The threshold for convergence was set to the default of 0.05. For identifiability only the total order sensitivity measure was used. The threshold for identifiability was varied to find the optimal value. The optimal value was determined by calculating accuracy, precision, recall and F1 [30] for each threshold value compared to the identifiability results from the NONMEM simulation and selecting the threshold with the highest F1 value. The highest F1 value indicates the optimal balance between precision and recall. The statistics were calculated as follows:

$$accuracy = \frac{true\ positives + true\ negatives}{true\ positives + true\ negatives + false\ positives + false\ negatives} \tag{39}$$

$$precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{40}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives} \tag{41}$$

$$F1 = 2\frac{precision * recall}{precision + recall} \tag{42}$$

$$\tag{43}$$

Positive here means that the parameter is unidentifiable and it follows that negative means that the parameter is identifiable. The reference here is the results from the NONMEM simulation. Thus, for example, a false positive is when a parameter is identifiable in the NONMEM simulation, but marked as unidentifiable by pksensi.

4.6 Fisher information Matrix identifiability analysis

The PopED package was also used to generate the Fisher Information Matrix in the same manner as the data sets were generated. The only difference is that now not only the ODEs of the full TMDD model were used, but those for the QSS model and the Michaelis-Menten model as well. This resulted in 63 different Fisher Information Matrices for each scenario used in the simulation approach. These matrices were then normalized for the structural model parameter values. The variability parameters are not normalized as these are already relative values. The matrix was normalized by matrix multiplication of the FIM with a diagonal with the true parameter values and the rest of the values set to 1. Finally, the diagonal was multiplied with the resulting matrix.
Next, the eigenvalues and vectors of the matrix were calculated. The eigenvalues indicate the size of effect on the output of a step in the direction of the eigenvector. A

eigenvalue of zero or close to zero indicates unidentifiability in that direction. Thus, to mark directions as unidentifiable a threshold on the eigenvalues needs to be determined, below which the direction is marked as unidentifiable. Two approaches were tested for selecting the unidentifiable directions, firstly all directions that have a eigenvalue lower than a certain threshold and secondly the largest jump between eigenvalues on the log scale was selected as the threshold for identifiability. The unidentifiable parameters are then selected by finding the largest absolute value in the eigenvector, which is "the main direction" of that eigenvector. The optimal threshold value was selected in a similar manner as for the variance-based sensitivity analysis, by varying the threshold and selecting the optimal F1 value.

4.7 Analyzing the results

The NONMEM estimated parameters were compared to the true (used in the simulations) parameter values. Firstly, the relative absolute differences between these parameter values were calculated for all the replicates as follows: $difference = \frac{|estimated\ parametervalue - true\ parameter\ value|}{true\ parameter\ value}$.
A difference of larger than 20% compared to the true value was assumed to indicate an unidentifiable parameter. The comparison to 0.2 (20%) is done by a sign test, as preliminary analysis showed the differences were neither (log-)normally distributed nor symmetric. The eFAST and the FIM method are then compared to the NONMEM results. As now for each experiment the identifiability of each parameter is decided for all methods, the accuracy, precision and F1 value of the eFAST and FIM method are calculated. To study the identifiability of the parameters in different sampling schemes the NONMEM results were studied in more detail by creating boxplots of the distribution of the estimated parameter values relative to the true parameter values. The width of the inter-quartile range is a measure of the precision of the parameter estimates. The difference between the median and the true value is a measure of accuracy.

5 Results

5.1 Simulation data

An example of a simulated data set created with PopED, according to the specifications in the methods can be seen in figure 5. It shows that the different dose groups cover a wide range of concentrations and the sampling is sufficiently dense and long enough to span across all phases of the typical TMDD concentration-time profile (See figure 3).

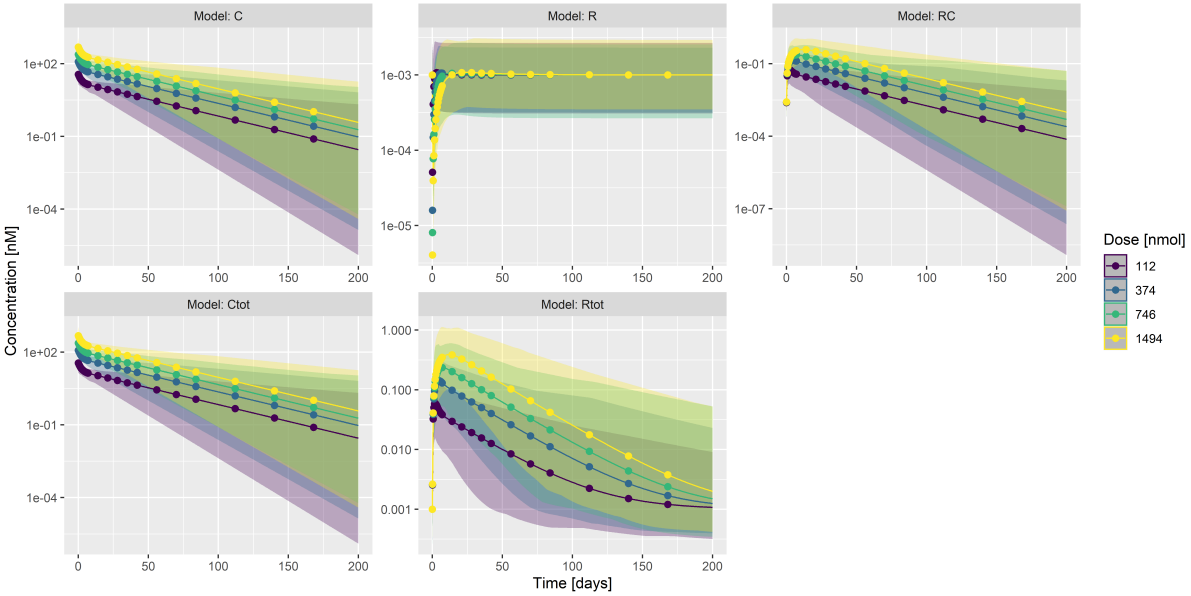

Figure 5: Concentration-time profiles of simulated data with TMDD model (k_{deg}=82 /day, k_{int}=0.01 /day). Sampling times are shown as dots and solid lines represent the typical profile. The shaded areas are the IIV. The profiles are coloured per dose group.

Figure 6 shows an overview of the concentration-time profiles for the different parameter value sets. It compares the profiles generated by the model simplifications (QSS and MM) to the TMDD profile. There is very little difference between the QSS and TMDD profile visible, except in the initial state. The initial drug-receptor complex concentration for QSS is at the value of $R0$ = 0.001 nM, while for TMDD it starts at 0 nM and then rapidly approaches the QSS profile. The difference of the TMDD profile with the MM approximation is far more pronounced and depends strongly on the parameter set. Firstly, in the MM approximation the total receptor concentration is constant. Additionally, the complex concentration does not initially increase/decrease depending on the ratio of k_{deg} and k_{int}. This is most noticeable when k_{deg}=82 /day and k_{int} is 0.01 /day, as this causes a strong increase of the complex concentration in TMDD and QSS.
The difference in k_{int}, also resulted in very different concentration-time profiles. There is a marked difference in the length of phase 2, which is longer when k_{int} is smaller. This also means that while for k_{int} = 0.01 /day there are samples in all phases, the last sample for k_{int} = 40 /day is taken in the third phase at the highest dose level.
If k_{int}=40 /day and $k_{deg} > k_{int}$, the total receptor concentration decreases from the steady state concentration initially. While the concentration increases if $k_{int} < k_{deg}$

before returning to the equilibrium.

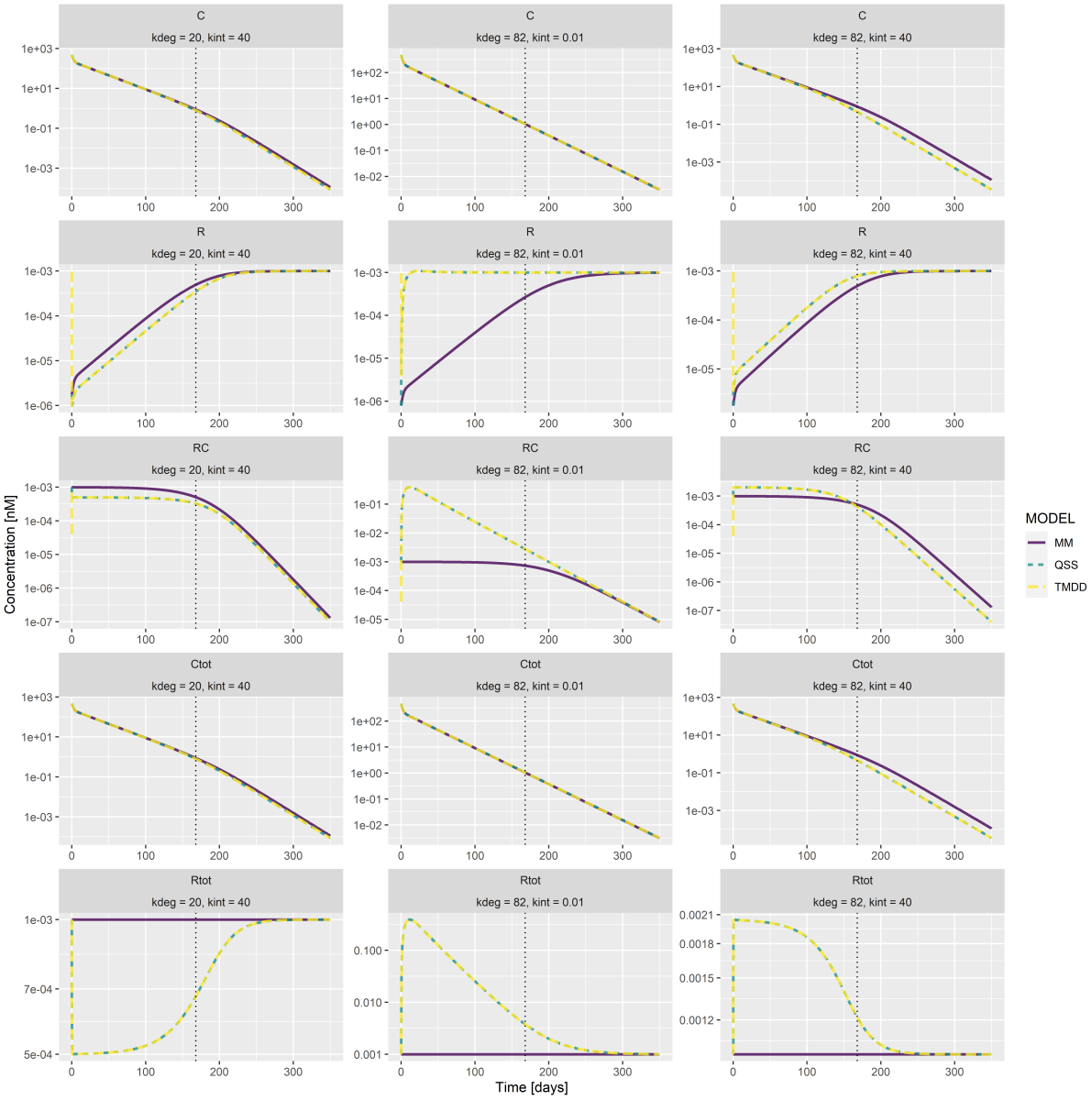

Figure 6: Simulations of concentration-time profiles for the different parameter sets and model approximations with dose is 1494 nmol. The vertical dotted black line indicates the final sampling point in the simulation data sets.

5.2 NONMEM parameter estimation

5.2.1 Initial results exploration

The settings of all experiments can be found in appendix 7.1. For each of the 63 experiments, 100 models with different initial estimates were constructed. All 6300 models were run with NONMEM. From the 6300 runs, 4639 minimized successfully. 1587 of the remaining runs crashed due to rounding errors. 74 runs crashed due to other reasons, mostly parameter estimates approaching zero. These 74 runs were

excluded from the results. Every experiment had at least 93 runs left. The number of excluded runs per experiment can be found in figure 7.
The decision was made to include the runs with rounding errors as the distribution over the different data sets and models was uneven (figure 8). Excluding runs with rounding errors may have resulted in skewed results, as the sign test for unidentifiability would have less power for data sets with less runs. Notably, the runs with the Michaelis-Menten approximation with a data set with only the biomarkers of free (C, dark blue) or total drug ($Ctot$, yellow) concentration are over-represented in the rounding errors. In addition, the distribution of the estimates of runs with rounding errors versus runs without, did also not seem to differ much on a visual inspection of the relative parameter estimates (figure 9).

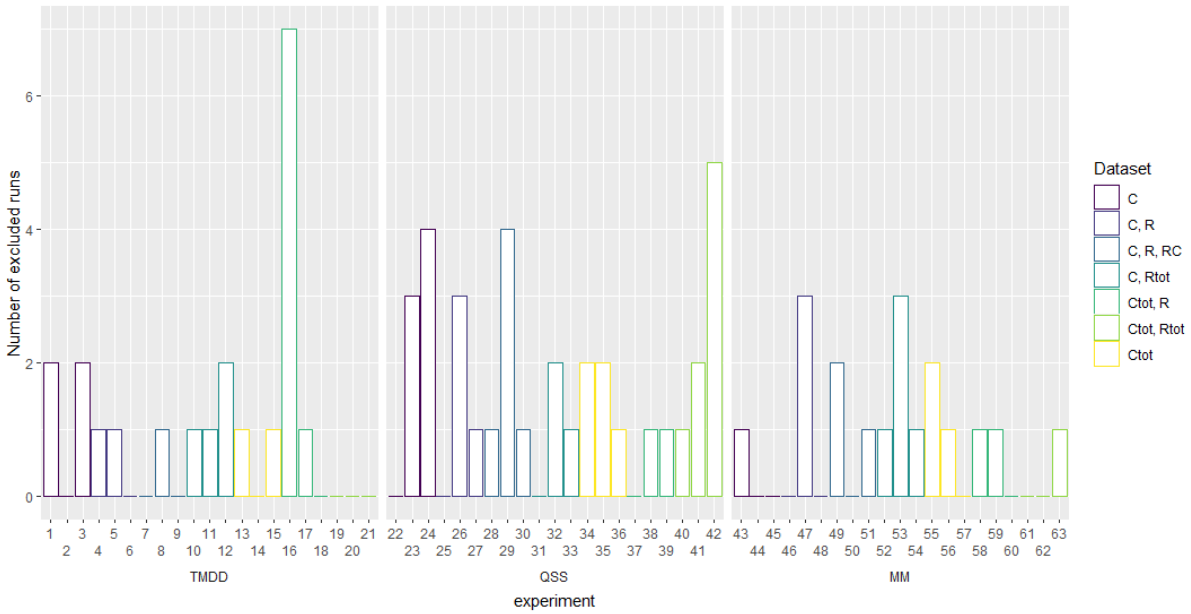

Figure 7: Number of excluded runs per experiment colored by the data set. Runs with rounding errors are not excluded.

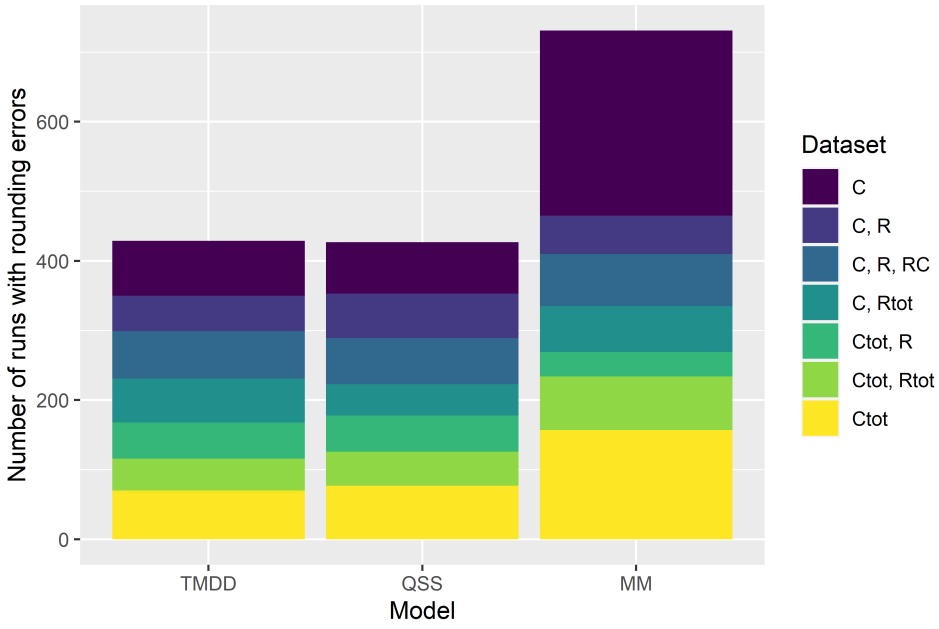

Figure 8: Number of runs with rounding errors per model colored by the data set.

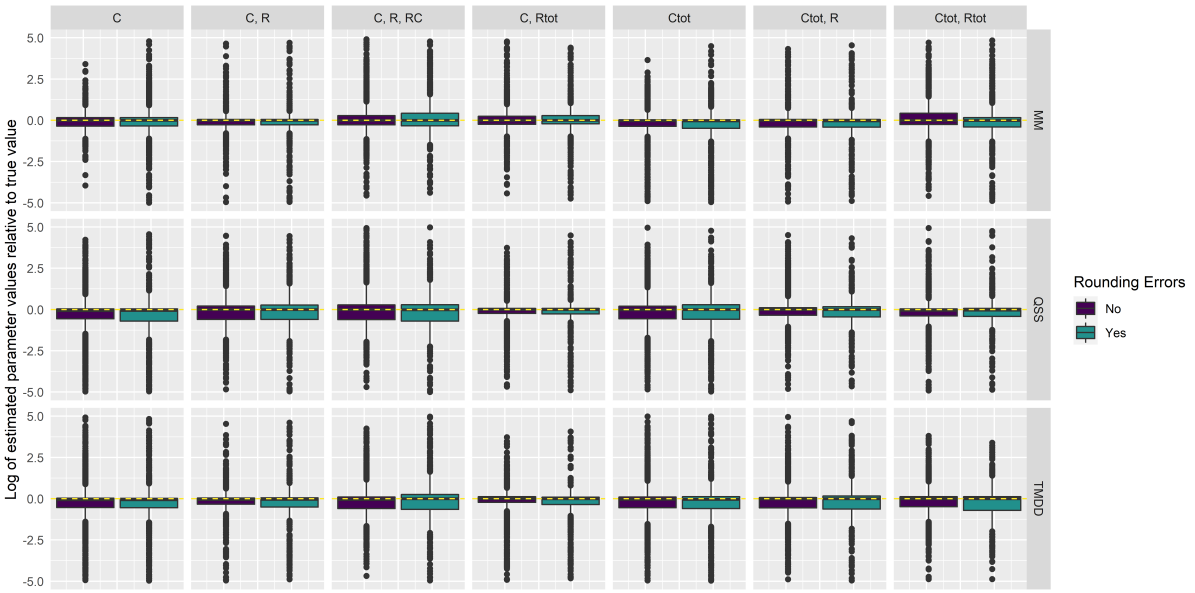

Figure 9: Boxplots of the distribution of natural logarithm of the relative estimated parameter values of all parameters over the data set and the model type, split by successful minimization or rounding errors. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [-5,5], so not all data may be included in the plot.

The absolute relative difference from the true parameter values were compared with a sign test to a value of 0.2 (20%), which was assumed to be the acceptable relative deviation from the true value. The number of parameters marked (un)identifiable by this test, can be found in table 4.

Table 4: Number of parameters marked as (un)identifiable by the sign test applied to the absolute relative differences of the NONMEM estimated values from the true value. The ratio between unidentifiable/identifiable parameters is provided as well. The results are provided for all parameters in all experiments and for only the fixed-effects parameters.

	Number of unidentifiable parameter values (True)	Number of identifiable parameter values (False)	Ratio
All parameters	121	752	0.16
Fixed-effects parameters	55	449	0.12

5.2.2 Results analysis

The parameter estimation for the different model types and parameter values relative to the true (simulation parameter) values can provide valuable information on the identifiability of parameters in TMDD models under different conditions. These results must be interpreted with caution because per experiment only one data set, with 24 individuals with parameter values randomly sampled from the specified distribution, was generated. A few of the more noteworthy results are discussed. Boxplots providing an overview of the results of the remaining parameter estimation can be found in appendix 7.2.
Firstly, figure 10 presents a breakdown of the results of the parameter estimation in NONMEM of the drug-target complex internalization parameter (k_{int}) for the different models and parameter sets. These and all the other boxplots presented hereafter show the distribution of the estimated parameter value relative to the true (simulated) parameter value. It is important to note that the plots have been limited to a factor 100 larger and smaller than the true parameter value for a more clear visualization of the bulk of the results.
What stands out in figure 10 is the wider range of predictions in the TMDD model with only the free or total drug concentration, compared to the other data sets for all three parameter sets. There also seems to be a strong ($\sim$10x) over-estimation in the value of k_{int} for the data set with free drug, free target and complex data for the QSS model with k_{int} = 0.01 /day compared to the parameter sets with k_{int} = 40 /day. However, it should be taken into account that these are relative values and as k_{int} is small, a small over-prediction results in a large relative over-prediction.
This is to a lesser degree also the case for the TMDD model with only free drug & free target concentration and total drug & total target concentration. While in contrast the experiments with only free drug & total receptor concentration and total drug and

total receptor concentration show an under-prediction in varying degrees for low values of k_{int}. Hence, the results for k_{int} indicate that it can be most accurately identified for higher values of k_{int} and when target data is available.

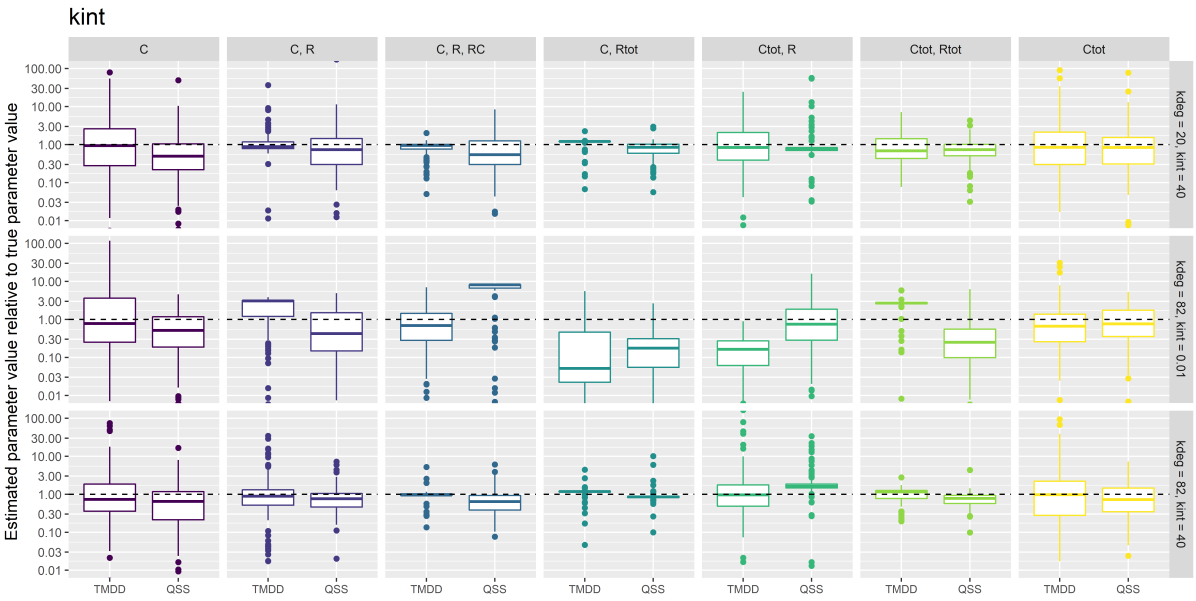

Figure 10: Plots of the estimated parameter value k_{int} (/ day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

The results for the rate of target degeneration (k_{deg}) are shown in figure 11. As with k_{int}, the range of estimated values for k_{deg} is wider when there is only free or total drug concentration in the data set. It is also apparent from this figure that addition of the drug-target complex data has little to no added value for the prediction of k_{deg} as the IQR is neither smaller nor more accurate than for the data sets which only contain the free drug concentration. Only for k_{int} = 0.01 and biomarkers C & R and $Ctot$ & R, is there a slight under-prediction in the value of k_{deg}. This might be caused by a bias introduced through the randomly sampled data set, as there is IIV on k_{deg}.

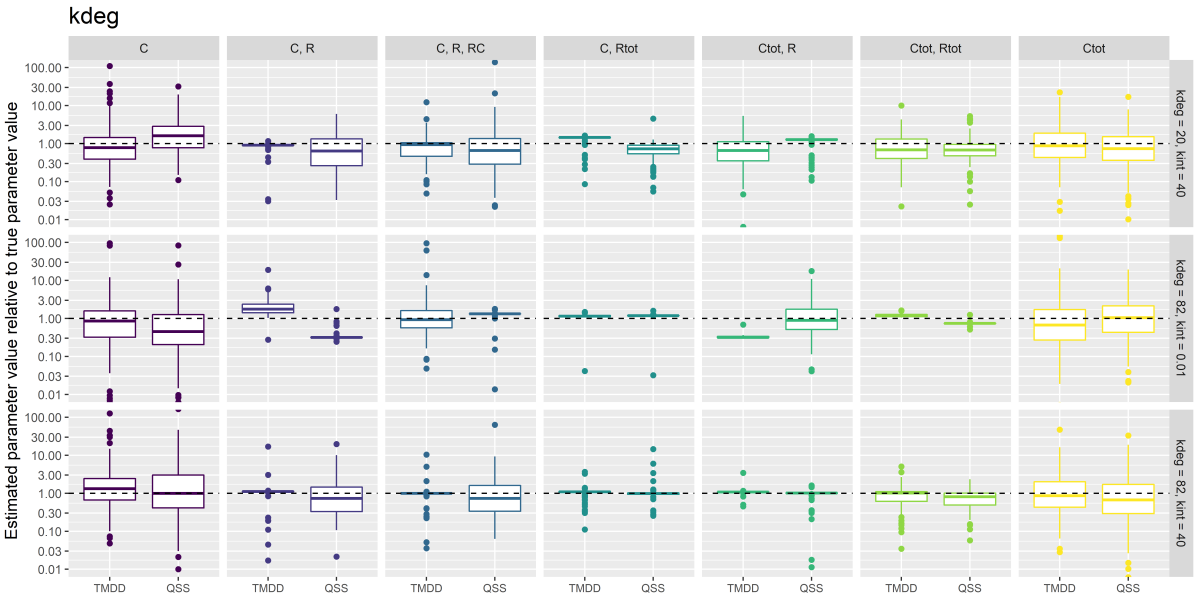

Figure 11: Plots of the estimated parameter value k_{deg} (/ day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

From figure 12 with the estimated parameter boxplots of the Michaelis-Menten constant (K_m) it can be seen that the prediction strongly depends on the parameter set and the data set. The most striking example of this is the extreme over-prediction of the value of K_m when k_{int} is small and the free receptor concentration without complex concentration is present in the dataset. This is better visible in figure 13, where the limit to the relative parameter estimation value is higher. Interestingly, the prediction is more accurate with only concentration data or total receptor concentration in the data set. This behavior may be explained by the difference in the free receptor and total receptor profile between MM and TMDD/QSS for low values of k_{int}, see figure 3).
Overall, these results show that the parameter identification depends on the biomarkers in the data set, the parameter values and the model used.

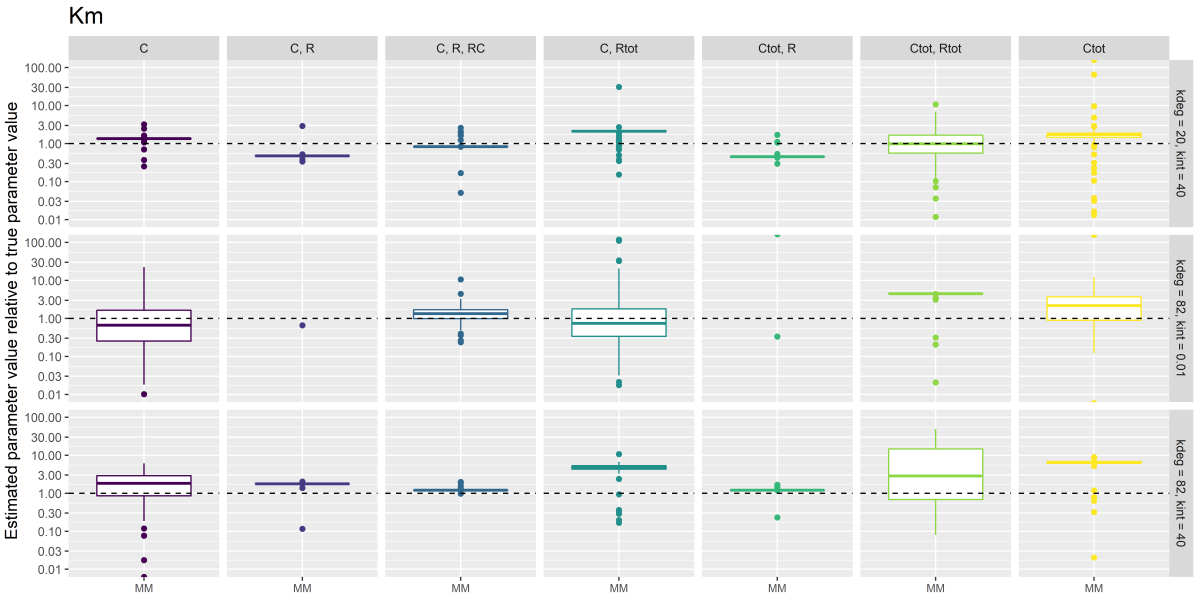

Figure 12: Plots of the estimated parameter value K_m (nM) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

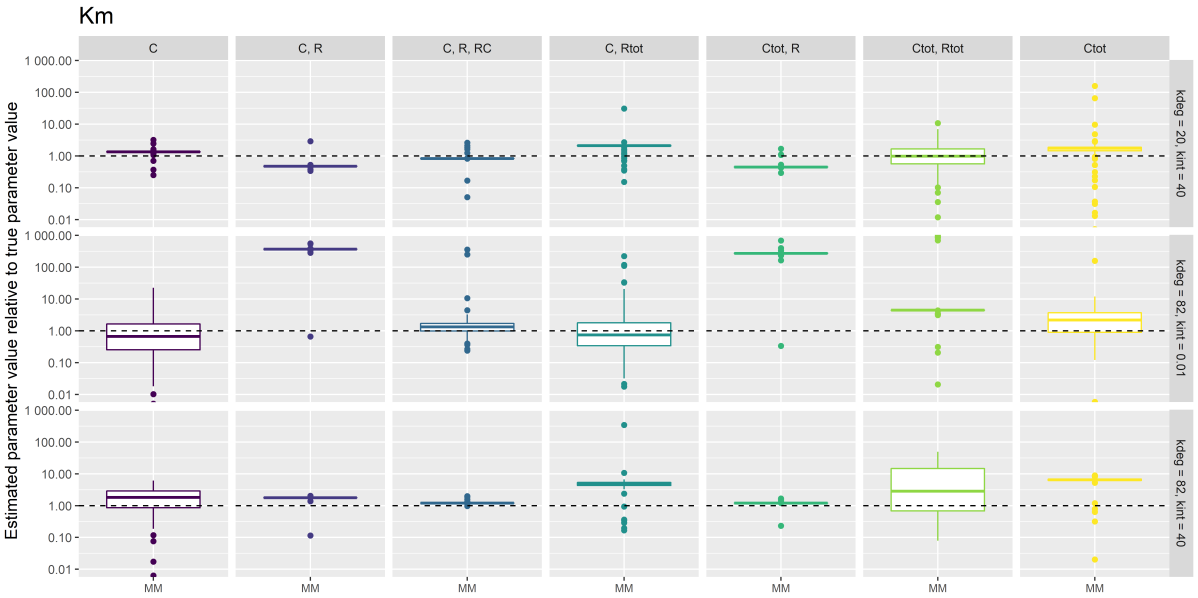

Figure 13: Plots of the estimated parameter value K_m (nM) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to 750, so not all data may be included in the plot.

5.3 Sobol Sensitivity Analysis

One of the two tested methods for predicting the identifiability of parameters in TMDD models is the variance-based sensitivity analysis eFAST for predicting Sobol sensitivity indices. Pksensi also provides a tool for monitoring the convergence of the sensitivity index (scale of 0-1). The default value for convergence is smaller than 0.05 on the convergence index, however Hsieh et al. use 0.1 as threshold in their study on applying global sensitivity analysis [13]. In figure 14 the convergence of the parameters over each experiment is plotted. It is apparent that many of the parameters have not reached convergence with a sample size of 10000, using 0.1 as threshold. Especially, for many of the experiments the clearance has not converged and overall the QSS model experiments (experiment number 21-42). Because the non-convergence was imbalanced over the different model types, the decision was made to include these predictions in the results. Consequently, the results for these experiments should be interpreted with caution.
The default threshold for identifiability in the pksensi package is 0.05 as well, but by varying the threshold value, it was found that 0.075 gave the highest F1 value and thus the best balance between precision and recall. However, the F1 value varies little over the entire range of threshold values, due to the precision being 25% over the entire range of boundary values. A threshold of 0.075 is used for the comparison to the NONMEM estimates. The accuracy at this threshold is 0.8, which means that 80% of the predictions is correct. However, there is a "class imbalance" as there are more parameters predicted to be identifiable (i.e False) than unidentifiable (i.e. True), thus the accuracy does not reflect the quality of the variance-based sensitivity analysis as a classifier for unidentifiable parameters well (See table 4). More informative is the precision. The precision at the selected threshold is 0.23, so 23% of the parameters that eFAST predicts are unidentifiable are unidentifiable according to the NONMEM results. Although, the precision is low, eFAST performs better than a random classifier where a precision of 12% is expected according to the class balance (see table 4).

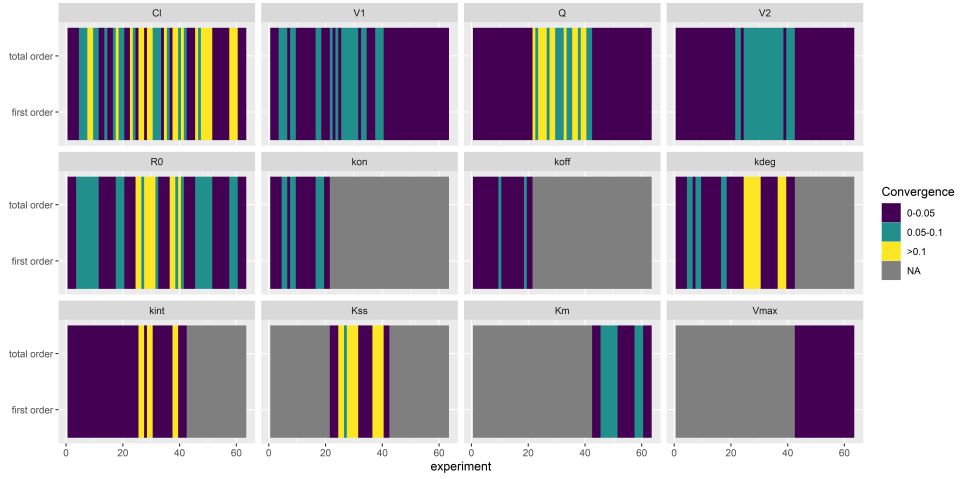

Figure 14: Plot of the first order and total order Sobol indices convergence for the different parameters. The colours indicate the degree of convergence.

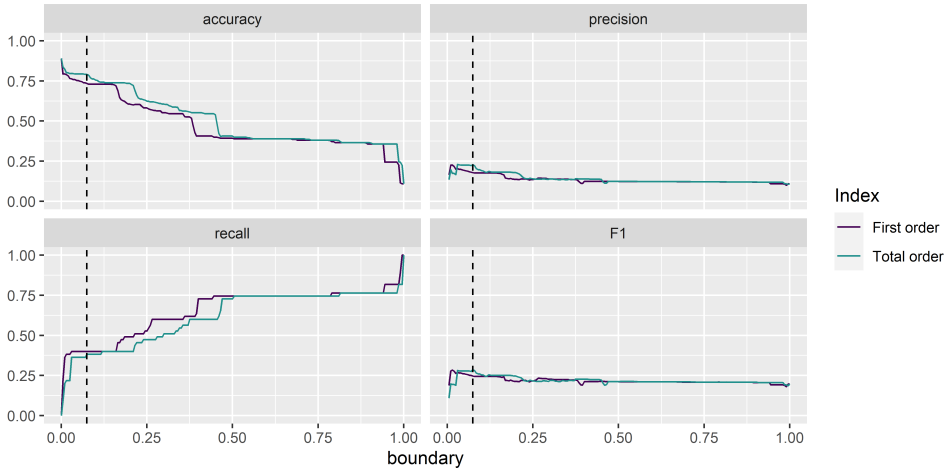

Figure 15: Plots of the accuracy, precision, recall and F1 of the eFAST variance-based sensitivity analysis with as reference the NONMEM simulation results. The Sobol indices range between 0 and 1. The purple line represents the first order sensitivity and the blue line represents the total order sensitivity. The dotted line indicates the selected threshold value (0.075), which has the highest F1 value.

5.4 Fisher Information Matrix

The Fisher information matrix can be used to determine the identifiability of fixed effects and random effects parameters. Low eigenvalues indicate an unidentifiable direction. To find the best threshold for the eigenvalues, different threshold values were tested, as well as using the largest difference between two eigenvalues on the log scale as the threshold. The eigenvalues were calculated relative to the sum of the eigenvalues from the FIM, so the range of values is between 0 and 1. The accuracy, precision, recall and F1 statistics of this experiment are plotted in figure 16. The F1 value did not vary much in the entire range of tested threshold values $(1 \cdot 10^{-10} - 0.01)$. The threshold with the highest F1 value $1 \cdot 10^{-8}$ was selected for comparison to the NONMEM estimated values, as this represents the optimal balance between the recall and precision. The recall for the studied threshold values varied between 0.55 and 0.85, which means that between 55% and 85% of unidentifiable parameters are found. However, the precision is lower than 0.25 over the entire range of threshold values, which indicates that only less than 25% of parameters marked as unidentifiable by the FIM method are actually unidentifiable (according to the simulations). This is around the same precision achieved by reached by the variance-based sensitivity method.
The ratio of unidentifiable to identifiable for all parameters is 0.16, thus the precision of the FIM method is better than a random classifier.

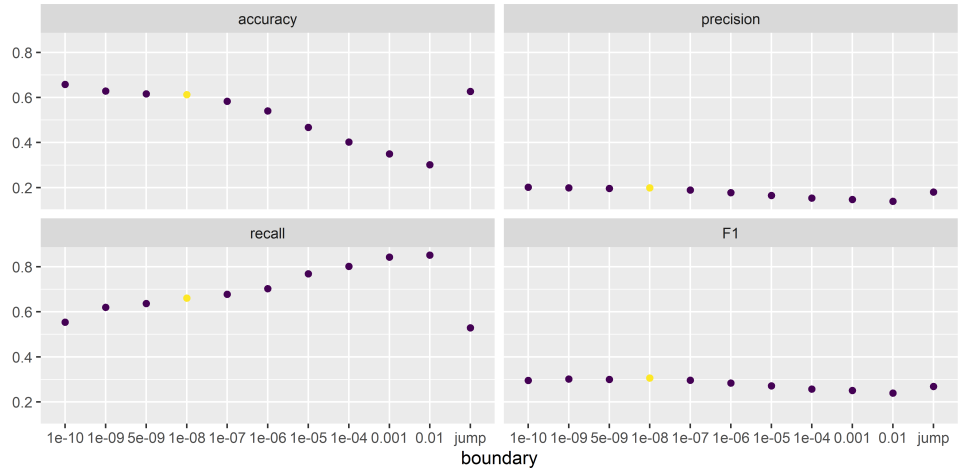

Figure 16: Plots of the accuracy, precision, recall and F1 of the Fisher Information matrix with as reference the NONMEM simulation results. The threshold values are the eigenvalues relative to the largest eigenvalue, thus the largest possible value is 1. The jump threshold is the threshold on identifiable eigenvalues set by selecting the largest difference between two eigenvalues on the log-scale. The yellow dot indicates the selected threshold value, which has the highest F1 value.

5.5 Comparing the methods

Using the selected boundaries for the eigenvalues from the FIM and the Sobol indexes calculated by eFAST the identifiability of each parameter was determined. In figure 17 these results compared to the results from the NONMEM estimation are presented. It is clear that the results from the FIM and Sobol indexes are somewhat in discordance with both each other and the NONMEM estimates. A positive result is that the predictions for Cl, $V1$, Q and $V2$ indicate that they are mostly identifiable over the experiments for all three methods. The eFAST method predicted that the kinetics parameters Cl, $V1$, Q and $V2$ are always identifiable and the FIM that k_{off}, k_{deg} and k_{int} are always identifiable. While the NONMEM simulation results indicated that for all the parameters (except $V2$) there are conditions where it is not identifiable. Interestingly, for V_{max} the eFAST indicates that is not identifiable in any experiment, while the FIM indicates that it is identifiable in every experiment, and the NONMEM results suggest it sometimes is, and sometimes is not. In contrast, the FIM predicted that for most experiments the k_{on}, k_{ss} and K_m are unidentifiable, while the eFAST method predicted more identifiable parameters here.

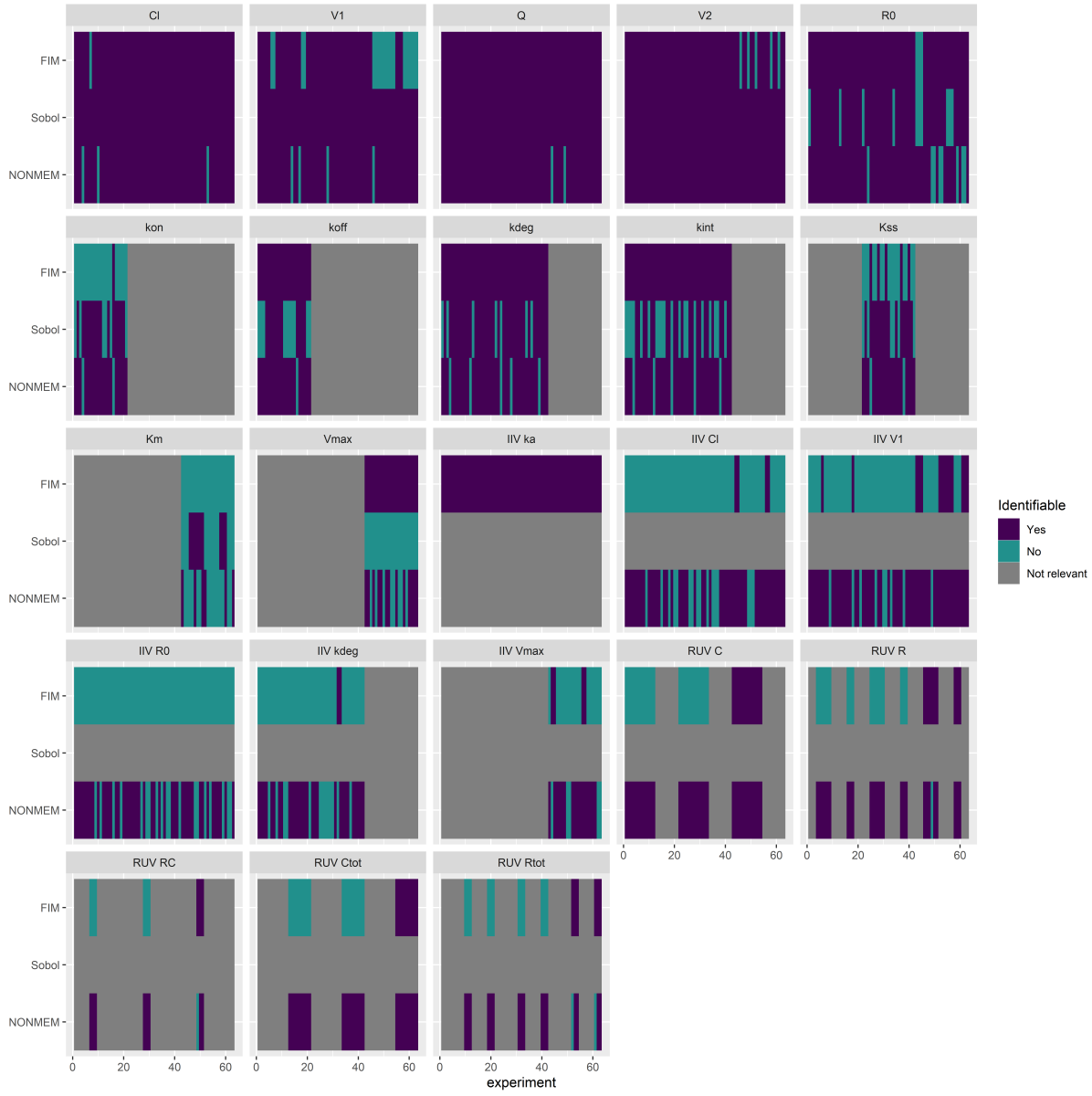

Figure 17: Plots of the accuracy, precision, recall and F1 of the Fisher Information matrix with as reference the NONMEM simulation results. The threshold values are the eigenvalues relative to the largest eigenvalue, thus the largest possible value is 1. The jump threshold is the threshold on identifiable eigenvalues set by selecting the largest difference between two eigenvalues on the log-scale. The yellow dot indicates the selected boundary value, which has the highest F1 value.

The statistics confirmed the visual indications. Both the eFAST and the FIM method resulted in comparable low precision scores, 0.23 and 0.20 respectively, although both were as mentioned higher than expected for a random classifier. The recall is considerably higher for the FIM (0.66) than the eFAST (0.38) method, with only a small reduction in precision. However, the eFAST method has a higher accuracy (0.79) than the FIM (0.58). This suggest that the eFAST has more true negative values, i.e. parameters correctly predicted to be identifiable, but less true positive values than the FIM. See table 5 for the statistics per parameter.

Table 5: Accuracy, precision, recall and F1 for the eFAST sensitivity method and the Fisher Information Matrix (FIM) method with the NONMEM estimates as reference. Missing values are caused by zero denominators in the calculations.

	parameter	method	accuracy	precision	recall	F1
3	Cl	eFAST	0.95	-	0.00	-
4	Cl	FIM	0.94	0.00	0.00	-
5	$V1$	eFAST	0.94	-	0.00	-
6	$V1$	FIM	0.67	0.05	0.25	0.09
7	Q	eFAST	0.97	-	0.00	-
8	Q	FIM	0.97	-	0.00	-
9	$V2$	eFAST	1.00	-	-	-
10	$V2$	FIM	0.92	0.00	-	-
11	$R0$	eFAST	0.71	0.00	0.00	-
12	$R0$	FIM	0.83	0.00	0.00	-
13	k_{on}	eFAST	0.62	0.00	0.00	-
14	k_{on}	FIM	0.05	0.05	0.50	0.09
15	k_{off}	eFAST	0.48	0.00	0.00	-
16	k_{off}	FIM	0.95	-	0.00	-
17	k_{deg}	eFAST	0.79	0.17	0.20	0.18
18	k_{deg}	FIM	0.88	-	0.00	-
19	k_{int}	eFAST	0.55	0.15	0.60	0.24
20	k_{int}	FIM	0.88	-	0.00	-
21	K_{ss}	eFAST	0.62	0.00	0.00	-
22	K_{ss}	FIM	0.24	0.06	0.50	0.11
23	K_m	eFAST	0.57	0.75	0.60	0.67
24	K_m	FIM	0.71	0.71	1.00	0.83
25	V_{max}	eFAST	0.38	0.38	1.00	0.55
26	V_{max}	FIM	0.62	-	0.00	-
30	$IIV\ Cl$	FIM	0.32	0.27	1.00	0.43
32	$IIV\ V1$	FIM	0.24	0.16	1.00	0.27
34	$IIV\ R0$	FIM	0.30	0.30	1.00	0.46
36	$IIV\ k_{deg}$	FIM	0.31	0.31	1.00	0.47

38	$IIV\ V_{max}$	FIM	0.43	0.27	0.80	0.40
40	$RUV\ C$	FIM	0.33	0.00	-	-
42	$RUV\ R$	FIM	0.30	0.00	0.00	-
44	$RUV\ RC$	FIM	0.11	0.11	1.00	0.20
46	$RUV\ C_{tot}$	FIM	0.33	0.00	-	-
48	$RUV\ R_{tot}$	FIM	0.22	0.00	0.00	-
49	all	eFAST	0.79	0.23	0.38	0.29
50	all	FIM	0.58	0.20	0.66	0.31

6 Discussion

The initial objective of this study was to determine the identifiability of TMDD model parameters under different conditions. To this end models were fitted on simulated data sets with different model simplifications, parameter values and biomarkers. In the results some of the more interesting findings from this analysis were highlighted.

Before these results are discussed, some remarks on the practical implementation of this study. Apart from the of course long runtime ($\sim$days) of such a large amount of models in NONMEM, several other challenges were encountered. To run models with different biomarkers and different initial values, different NONMEM models need to be created. As 6300 models is too much to code manually, three templates for these files were generated, one for each model type (TMDD, QSS and MM) in R using the package whisker [31]. Another issue was that the ODE for the three model types had to be declared in three different languages, for NONMEM, the PopED package and pksensi. All three model declarations were used for simulating and these simulations were visually compared for ensuring integrity of the methods. For future work it may be beneficial to automatize this process. The entire process of creating experiments, generating data sets, generating the NONMEM model files, calculating the FIM's and running the eFAST sensitivity analysis was coordinated from R. The initial intention was to run all these processes in a single for-loop and start with a random seed for reproducibility. However, due to the heavy workload of the sensitivity analysis R crashes repeatedly during this process. Requiring a restart part-way through the process, means that the random seed will be reset each time, and reproducibility is thus lost. In a further study, these processes should be separated to avoid this problem.

Turning now to the results of the NONMEM simulations, in line with the expectations it was found that when there is only free or total drug data, the distribution of parameter estimates for dynamics parameters (k_{int}, k_{deg} and K_m) is generally wider as there is no direct information on the target concentration. The most remarkable finding was the strong over-estimation for K_m, when there was free/total target concentration in the data set and k_{int} = 0.01 /day for the MM approximation. The estimation of k_{int} was actually better when there was only drug-concentration data in the data set. This is likely caused by the disagreement between the concentration-time profile of the target for the full TMDD model and the MM approximation with a low k_{int} value. A higher k_{int} value will shift the curve of the MM profile to the left, which agrees more with the receptor curve, while only influencing the drug-concentration profile on a limited scale. Slower drug-target complex degradation with values in the range of free-drug elimination is associated with soluble targets [32]. Therefore, it is important to take this into account when developing a model for a drug with a soluble target.

The disagreement between the different simplifications and its consequences evoke another point of discussion, namely the definition of identifiability. The definition of practical parameter identifiability in the introduction was stated as follows: parameter values can be accurately estimated given a finite amount of noisy

experimental data. If the model is not applicable to the data (as is the case here with MM models for low values of k_{kint}), according to the current definition this model is not identifiable. However, the question is if these are not two separate issues: identifiability and applicability. In this case another definition for identifiability might be more appropriate, e.g. parameter values can be uniquely estimated given a finite amount of noisy experimental data. With this definition no assumptions on the correctness of these parameter values are made. Then, in the analysis of the NONMEM parameter estimates the only criterion for the identifiability would be the width of the distribution of parameter estimates.

The analysis of the NONMEM estimates had several limitations and issues. Firstly, due to limited time and resources only a fraction of the possible situations was studied. The choice was made to only look at a single dose "phase 1" sampling and dosing scheme. However, the analysis could be expanded on many fronts, such as multiple dosing, different methods of dose administration, different sampling schemes, more individuals and other parameter sets or model simplifications. Secondly, while the initial parameter values in NONMEM were randomly sampled for the 100 replications, only one data set was used for each experiment. Because this is a mixed-effects model, the data values are selected with randomness due to the IIV and the RUV on the fixed effect parameters. For more reliable results, there should also be replications with re-sampled data sets. Important to note is that because per experiment only one data set, with 24 individuals with randomly sampled parameter values from the specified distribution, was generated. This may have resulted in the "true parameter value" not completely agreeing with the average parameter values for the individuals in the generated data set. This made the accuracy of the estimated parameter values difficult to assess and might have even resulted in incorrectly marking parameters as unidentifiable by the sign test. Furthermore, not all experiments had the same amount of runs due to a number of crashed runs. These experiments have less power in the sign test as there are less samples and this might have caused parameters to be classified as identifiable when they were not identifiable. Although, it is unclear to what extent these crashes have influenced the results.

The second aim of this study was to test a part of the proposed workflow (figure 4) for determining parameter identifiability. First, the results of Fisher Information Matrix method were compared to identifiability of parameters found by fitting the simulated data sets with NONMEM. The distribution of parameter estimates by NONMEM was transformed to a yes/no for identifiability by means of a sign test with 0.2. For the FIM method a threshold was set on the eigenvalues. The highest value of the corresponding eigenvector was then selected as the unidentifiable parameter. The threshold was varied to find the optimal balance between recall and precision through the F1 value. However, the precision was low (< 0.25) for the entire range of threshold values, meaning that less than 25% of the parameters marked as unidentifiable are actually unidentifiable. The recall was low as well, only 38% of the true unidentifiable parameters were marked as unidentifiable. This indicates that this approach is not suitable for determining the identifiability. However, the precision is better than that of a random classifier and with some adaptations it may be improved. It is possible that selecting just the highest value in the eigenvector is not sufficient and looking at

parameter combinations would give a higher precision. Another method could be to use the FIM generated by PopED in the method described by Eisenberg & Hayashi in [10] where identifiable parameter combinations are identified by a combination of the covariance-matrix from the FIM and log-likelihood profiling could be used instead of the method described here.

The other method that was tested is variance-based sensitivity analysis. A quantitative measure of sensitivity, the Sobol indices, were calculated using eFAST. The convergences of the indices can be checked with similar indices to the Sobol indices with the R package pksensi. These indices showed that many parameters had not converged with a sample size of 10000. In an analysis of convergence of variance-based sensitivity [33], it was found that sometimes sample sizes of more than 30,000 are necessary to reach convergence. However, they also suggested that one could also look at convergence of the screening, i.e. the split between identifiable and unidentifiable parameters, instead of the convergence of the sensitivity indices. The screening convergence might be reached with a smaller sampling size and is sufficient for the aims of the analysis in this study, thus this may be a better approach for future work. As with the FIM, the precision was low (max 23%) regardless of the threshold value on convergence. The recall was better than for the FIM, with a recall of 66%. However, the results may improve if convergence is reached for all parameters. Regarding the range of parameter values used, the original intent was to use the entire physiological range of parameter values in accordance with global parameter sensitivity analysis. However, initial analysis using such a large parameter range showed that the interaction term of the total index would then approach 1, making the results uninformative. In addition, convergence was very difficult to reach with such large ranges. Alternatively, the approach suggested in [13] was to select the parameter value ranges by preliminary simulations and choosing the range so that the simulated range covers the experimental data range.

For both the FIM method and the eFAST method, a threshold on sensitivity was set by selecting the value with the highest F1 value. The F1 value is the harmonic mean of the recall and precision. However, it may be better to weigh the recall and precision differently, as it may be more beneficial to simplify only when really necessary (so more "weight" on the precision) or to always simplify when there is any doubt about the identifiability (more "weight" on the recall). This is something that should be carefully considered and may also depend on the application of the model. Another point of deliberation is that both the FIM and eFAST method, only use the sampling points and the FIM an initial parameter estimate. To calculated the sensitivity they use the ODE's of the model they're testing the sensitivity of. However, as no actual data is used in these methods, if the model does not match the data the parameters will not be found to be unidentifiable. If this is a disadvantage or not depends on the chosen definition of identifiability. Regarding the practical applicability of these methods, it is important to note that the calculation of the FIM was very fast ($\sim$few minutes for all 63 experiments), in contrast to eFAST method which took over 24 hours to run. For a sampling size of 10000 with 10 replicates and four dose groups, a single run was around 20 minutes. However, as a larger sampling size may be needed the necessary time will increases accordingly. Moreover, the final step of the proposed workflow, the log-likelihood profiling, which was not tested yet, does use the

measured data point. This approach may capture unidentifiability missed by the FIM and eFAST, however will increases the time necessary for completing the workflow.

The final goal of this study was to identify "difficult" parameters, of which the identifiability was often predicted incorrectly by the FIM and eFAST method. However, as neither of these methods was successful, the results per parameter provided in table 5 were not informative. A more extensive investigation in which conditions the FIM and eFAST agree with the NONMEM results is necessary. This may also guide improvements of the proposed workflow.

In conclusion, the most important findings from the simulations were the overestimation of the k_{int} for low values of k_{int} with a Michaelis-Menten model using free/total target as biomarker in the dataset. This should be taken into account when developing models for soluble drug targets. Finally, the results for the tests of the steps of the proposed workflow were not encouraging. However, further research and possible adjustments are needed to make any definitive conclusions on the usefulness of this workflow for determining parameter identifiability for TMDD models.

References

[1] LB Sheiner, B Rosenberg, and VV Marathe. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of pharmacokinetics and biopharmaceutics*, 5(5):445—479, October 1977. ISSN 0090-466X. doi: 10.1007/bf01061728. URL https://doi.org/10.1007/bf01061728.

[2] Ss Beal, L. Boeckmann, Rj Bauer, and Lb Sheiner. Nonmem user's guides. (1989–2021). 2021.

[3] Gerhard Levy. Pharmacologic target-mediated drug disposition. *Clinical Pharmacology & Therapeutics*, 56(3):248–252, 1994. doi: https://doi.org/10.1038/clpt.1994.134. URL https://ascpt.onlinelibrary.wiley.com/doi/abs/10.1038/clpt.1994.134.

[4] Guohua An. Concept of pharmacologic target-mediated drug disposition in large-molecule and small-molecule compounds. *The Journal of Clinical Pharmacology*, 60(2):149–163, 2020. doi: https://doi.org/10.1002/jcph.1545. URL https://accp1.onlinelibrary.wiley.com/doi/abs/10.1002/jcph.1545.

[5] Donald Mager and William Jusko. General pharmacokinetic model for drugs exhibiting target-mediated drug disposition. *Journal of Pharmacokinetics and Pharmacodynamics*, 28:507–532, 12 2001. doi: 10.1023/A:1014414520282.

[6] L. Gibiansky, E. Gibiansky, T. Kakkar, and Peiming Ma. Approximations of the target-mediated drug disposition model and identifiability of model parameters. *Journal of Pharmacokinetics and Pharmacodynamics*, 35:573–591, 2008.

[7] Lambertus A Peletier and Johan Gabrielsson. Dynamics of target-mediated drug disposition: characteristic profiles and parameter identification. *Journal of pharmacokinetics and pharmacodynamics*, 39(5):429—451, October 2012. ISSN 1567-567X. doi: 10.1007/s10928-012-9260-6. URL https://europepmc.org/articles/PMC3446204.

[8] Donald Mager and Wojciech Krzyzanski. Quasi-equilibrium pharmacokinetic model for drugs exhibiting target-mediated drug disposition. *Pharmaceutical research*, 22:1589–96, 11 2005. doi: 10.1007/s11095-005-6650-0.

[9] Leonid Gibiansky and Ekaterina Gibiansky. Target-mediated drug disposition: New derivation of the michaelis-menten model, and why it is often sufficient for description of drugs with tmdd, 2010. URL https://www.page-meeting.org/?abstract=1728.

[10] Marisa C. Eisenberg and Michael A.L. Hayashi. Determining identifiable parameter combinations using subset profiling. *Mathematical Biosciences*, 256: 116–126, 2014. ISSN 0025-5564. doi: https://doi.org/10.1016/j.mbs.2014.08.008. URL https://www.sciencedirect.com/science/article/pii/S0025556414001631.

[11] Broeker Astrid and Sebastian G. Wicha. Assessing parameter uncertainty in small-n pharmacometric analyses: value of the log-likelihood profiling-based sampling importance resampling (llp-sir) technique. *Journal of Pharmacokinetics and Pharmacodynamics*, 47(3):219–228, 06 2020. URL https://www-proquest-com.ezproxy.leidenuniv.nl/scholarly-journals/assessing-parameter-uncertainty-small-n/docview/2412227281/se-2?accountid=12045.

[12] X-Y Zhang, MN Trame, LJ Lesko, and S Schmidt. Sobol sensitivity analysis: A tool to guide the development and evaluation of systems pharmacology models. *CPT: Pharmacometrics & Systems Pharmacology*, 4(2):69–79, 2015. doi: https://doi.org/10.1002/psp4.6. URL https://ascpt.onlinelibrary.wiley.com/doi/abs/10.1002/psp4.6.

[13] Nan-Hung Hsieh, Brad Reisfeld, Frederic Y. Bois, and Weihsueh A. Chiu. Applying a global sensitivity analysis workflow to improve the computational efficiencies in physiologically-based pharmacokinetic modeling. *Frontiers in Pharmacology*, 9:588, 2018. ISSN 1663-9812. doi: 10.3389/fphar.2018.00588. URL https://www.frontiersin.org/article/10.3389/fphar.2018.00588.

[14] A. Saltelli, S. Tarantola, and K. P.-S. Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1): 39–56, 1999. ISSN 00401706. URL http://www.jstor.org/stable/1270993.

[15] I.M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiments*, 4:407–4014, 1993.

[16] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996. ISSN 0951-8320. doi: https://doi.org/10.1016/0951-8320(96)00002-6. URL https://www.sciencedirect.com/science/article/pii/0951832096000026.

[17] KR Godfrey and WR Fitch. The deterministic identifiability of nonlinear pharmacokinetic models. *Journal of pharmacokinetics and biopharmaceutics*, 12(2):177—191, April 1984. ISSN 0090-466X. doi: 10.1007/bf01059277. URL https://doi.org/10.1007/bf01059277.

[18] Anne-Gaëlle Dosne, Martin Bergstrand, Kajsa Harling, and Mats O Karlsson. Improving the estimation of parameter uncertainty distributions in nonlinear mixed effects models using sampling importance resampling. *Journal of Pharmacokinetics and Pharmacodynamics*, 43(6):583–596, 2016. ISSN 1567-567X. doi: 10.1007/s10928-016-9487-8.

[19] Alexander P Browning, David J Warne, Kevin Burrage, Ruth E Baker, and Matthew J Simpson. Identifiability analysis for stochastic differential equation models in systems biology. *bioRxiv*, 2020. doi: 10.1101/2020.08.10.245233. URL https://www.biorxiv.org/content/early/2020/10/15/2020.08.10.245233.

[20] Hoai-Thu Thai, France Mentré, Nicholas H. G Holford, Christine Veyrat-Follet, and Emmanuelle Comets. Evaluation of bootstrap methods for estimating

uncertainty of parameters in nonlinear mixed-effects models: a simulation study in population pharmacokinetics. *Journal of pharmacokinetics and pharmacodynamics*, 41(1):15–33, 2014. ISSN 1567-567X.

[21] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25 15: 1923–9, 2009.

[22] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA, 2020. URL http://www.rstudio.com/.

[23] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

[24] Lars Lindbom, Pontus Pihlgren, Eirikur Jonsson, and Niclas Jonsson. Psn-toolkit - a collection of computer intensive statistical methods for non-linear mixed effect modeling using nonmem. *Computer methods and programs in biomedicine*, 79: 241–57, 10 2005. doi: 10.1016/j.cmpb.2005.04.005.

[25] Marco Foracchia, Andrew C Hooker, Paolo Vicini, and A Ruggeri. Poped, a software for optimal experiment design in population kinetics. *Computer Methods and Programs in Biomedicine*, 74, 2004.

[26] Joakim Nyberg, Sebastian Ueckert, Eric A Stroemberg, Stefanie Hennig, Mats O Karlsson, and Andrew C Hooker. Poped: An extended, parallelized, nonlinear mixed effects models optimal design tool. *Computer Methods and Programs in Biomedicine*, 108, 2012.

[27] Abhinav Tiwari, M. Kasaian, A. Heatherington, H. Jones, and F. Hua. A mechanistic pk/pd model for two anti-il13 antibodies explains the difference in total il-13 accumulation observed in clinical studies. *mAbs*, 8:983 – 990, 2016.

[28] Anti-human il13 therapeutic antibody (anrukinzumab). URL https://www.creativebiolabs.net/Anti-Human-IL13-Therapeutic-Antibody-anrukinzumab-13625.htm.

[29] Nan-Hung Hsieh, Brad Reisfeld, and Weihsueh A. Chiu. pksensi: An r package to apply global sensitivity analysis in physiologically based kinetic modeling. *SoftwareX*, 12:100609, 2020. ISSN 2352-7110. doi: https://doi.org/10.1016/j.softx.2020.100609. URL https://www.sciencedirect.com/science/article/pii/S2352711020303228.

[30] Yutaka Sasaki. The truth of the f-measure. *Teach Tutor Mater*, 01 2007.

[31] Edwin de Jonge. Whisker. URL https://github.com/edwindj/whisker.

[32] Leonid Gibiansky and Ekaterina Gibiansky. Target-mediated drug disposition model for drugs that bind to more than one target. *Journal of pharmacokinetics and pharmacodynamics*, 37:323–46, 08 2010. doi: 10.1007/s10928-010-9163-3.

[33] Fanny Sarrazin, Francesca Pianosi, and Thorsten Wagener. Global sensitivity analysis of environmental models: Convergence and validation. *Environmental Modelling & Software*, 79:135–152, 2016. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2016.02.005. URL https://www.sciencedirect.com/science/article/pii/S1364815216300251.

7 Appendices

7.1 Experiment Settings

Table 6: Settings for each experiment on the model type, the biomarkers in the dataset and the parameter values of k_{deg} and k_{int}.

Experiment	Model	Data set	Parameter set: k_{deg} (/day); k_{int} (/day)
1	TMDD	C	82; 0.01
2	TMDD	C	82; 40
3	TMDD	C	20; 40
4	TMDD	C, R	82; 0.01
5	TMDD	C, R	82; 40
6	TMDD	C, R	20; 40
7	TMDD	C, R, RC	82; 0.01
8	TMDD	C, R, RC	82; 40
9	TMDD	C, R, RC	20; 40
10	TMDD	C, R_{tot}	82; 0.01
11	TMDD	C, R_{tot}	82; 40
12	TMDD	C, R_{tot}	20; 40
13	TMDD	C_{tot}	82; 0.01
14	TMDD	C_{tot}	82; 40
15	TMDD	C_{tot}	20; 40
16	TMDD	C_{tot}, R	82; 0.01
17	TMDD	C_{tot}, R	82; 40
18	TMDD	C_{tot}, R	20; 40
19	TMDD	C_{tot}, R_{tot}	82; 0.01
20	TMDD	C_{tot}, R_{tot}	82; 40
21	TMDD	C_{tot}, R_{tot}	20; 40
22	QSS	C	82; 0.01
23	QSS	C	82; 40
24	QSS	C	20; 40
25	QSS	C, R	82; 0.01
26	QSS	C, R	82; 40

27	QSS	C, R	20; 40
28	QSS	C, R, RC	82; 0.01
29	QSS	C, R, RC	82; 40
30	QSS	C, R, RC	20; 40
31	QSS	C, R_{tot}	82; 0.01
32	QSS	C, R_{tot}	82; 40
33	QSS	C, R_{tot}	20; 40
34	QSS	C_{tot}	82; 0.01
35	QSS	C_{tot}	82; 40
36	QSS	C_{tot}	20; 40
37	QSS	C_{tot}, R	82; 0.01
38	QSS	C_{tot}, R	82; 40
39	QSS	C_{tot}, R	20; 40
40	QSS	C_{tot}, R_{tot}	82; 0.01
41	QSS	C_{tot}, R_{tot}	82; 40
42	QSS	C_{tot}, R_{tot}	20; 40
43	MM	C	82; 0.01
44	MM	C	82; 40
45	MM	C	20; 40
46	MM	C, R	82; 0.01
47	MM	C, R	82; 40
48	MM	C, R	20; 40
49	MM	C, R, RC	82; 0.01
50	MM	C, R, RC	82; 40
51	MM	C, R, RC	20; 40
52	MM	C, R_{tot}	82; 0.01
53	MM	C, R_{tot}	82; 40
54	MM	C, R_{tot}	20; 40
55	MM	C_{tot}	82; 0.01
56	MM	C_{tot}	82; 40
57	MM	C_{tot}	20; 40
58	MM	C_{tot}, R	82; 0.01

59	MM	C_{tot}, R	82; 40
60	MM	C_{tot}, R	20; 40
61	MM	C_{tot}, R_{tot}	82; 0.01
62	MM	C_{tot}, R_{tot}	82; 40
63	MM	C_{tot}, R_{tot}	20; 40

7.2 NONMEM parameter estimation

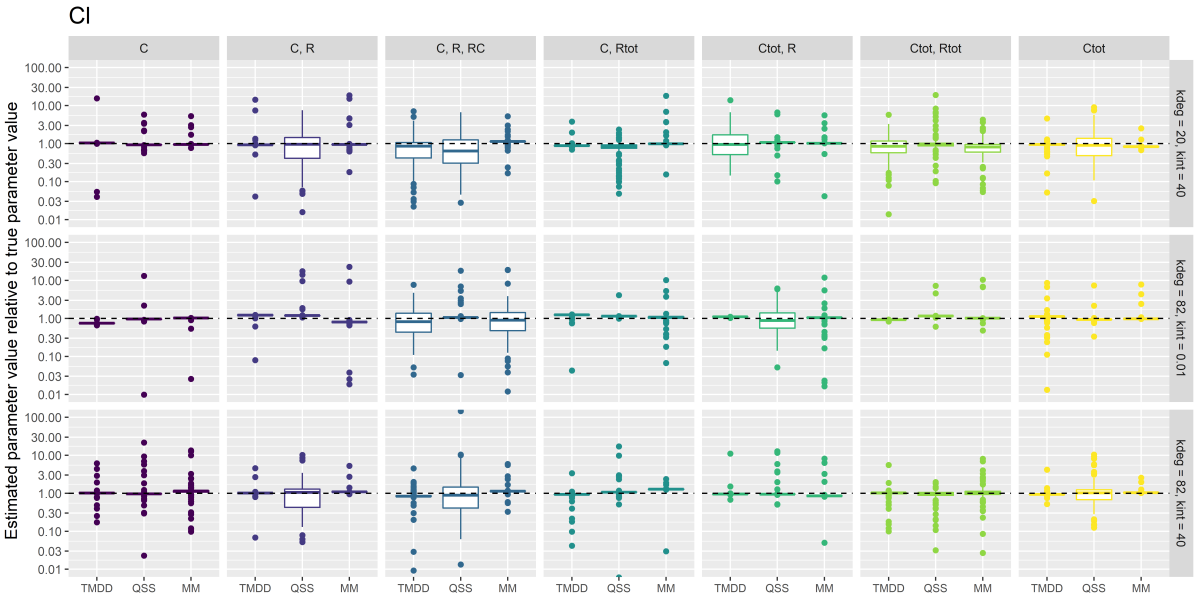

Figure 18: Plots of the estimated parameter value Cl (L/day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

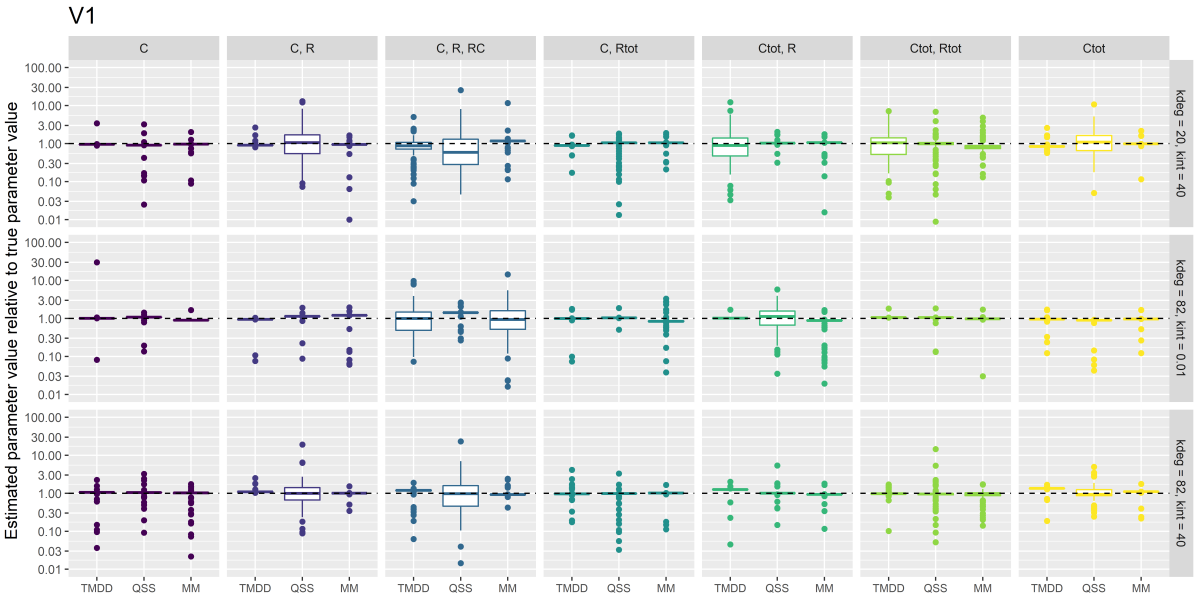

Figure 19: Plots of the estimated parameter value $V1$ (L) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

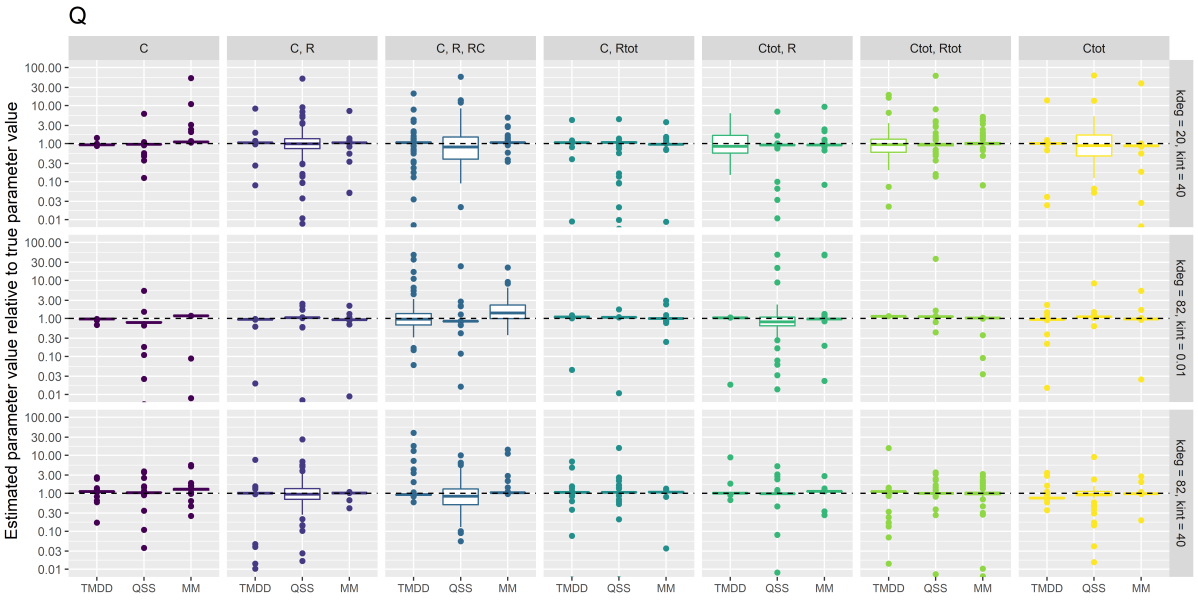

Figure 20: Plots of the estimated parameter value Q (L/day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

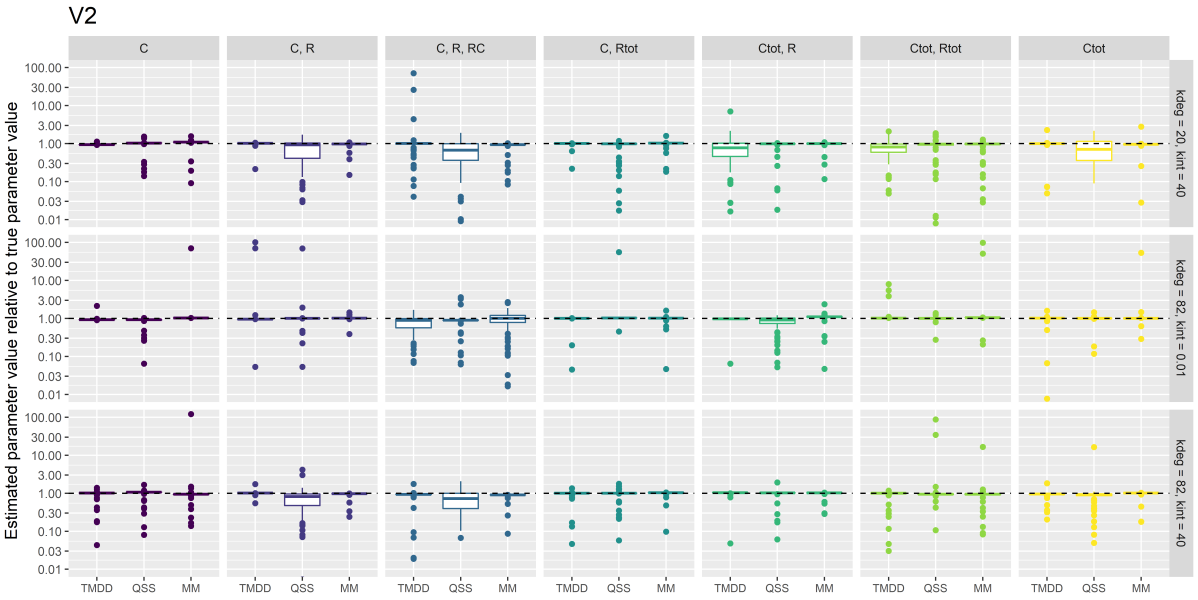

Figure 21: Plots of the estimated parameter value $V2$ (L) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

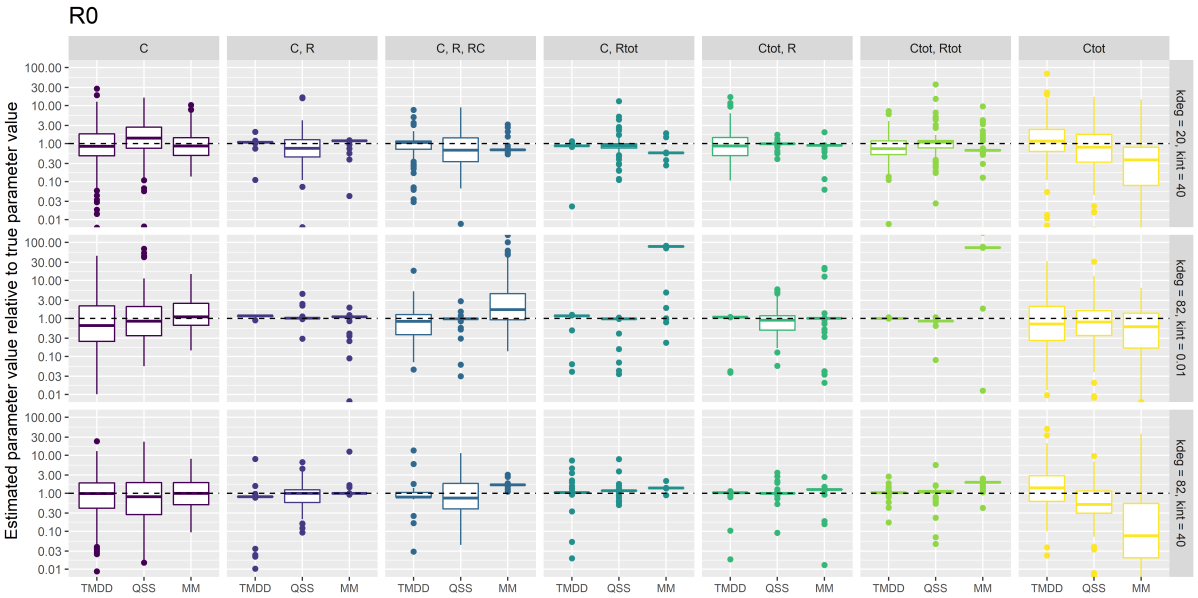

Figure 22: Plots of the estimated parameter value $R0$ (nM) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

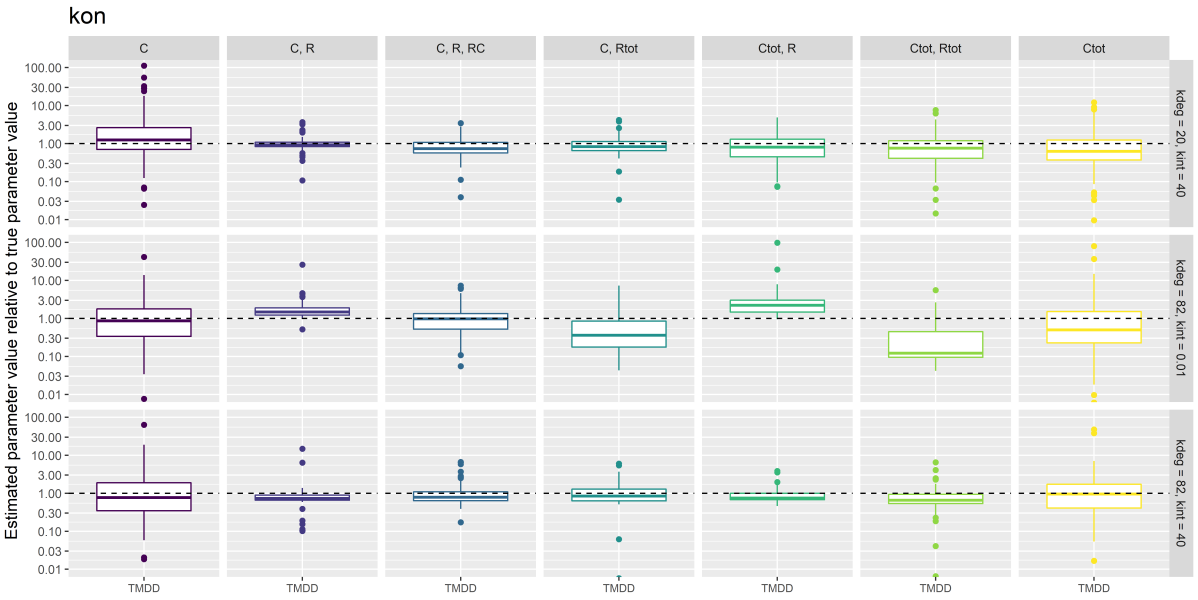

Figure 23: Plots of the estimated parameter value k_{on} (/nM/day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

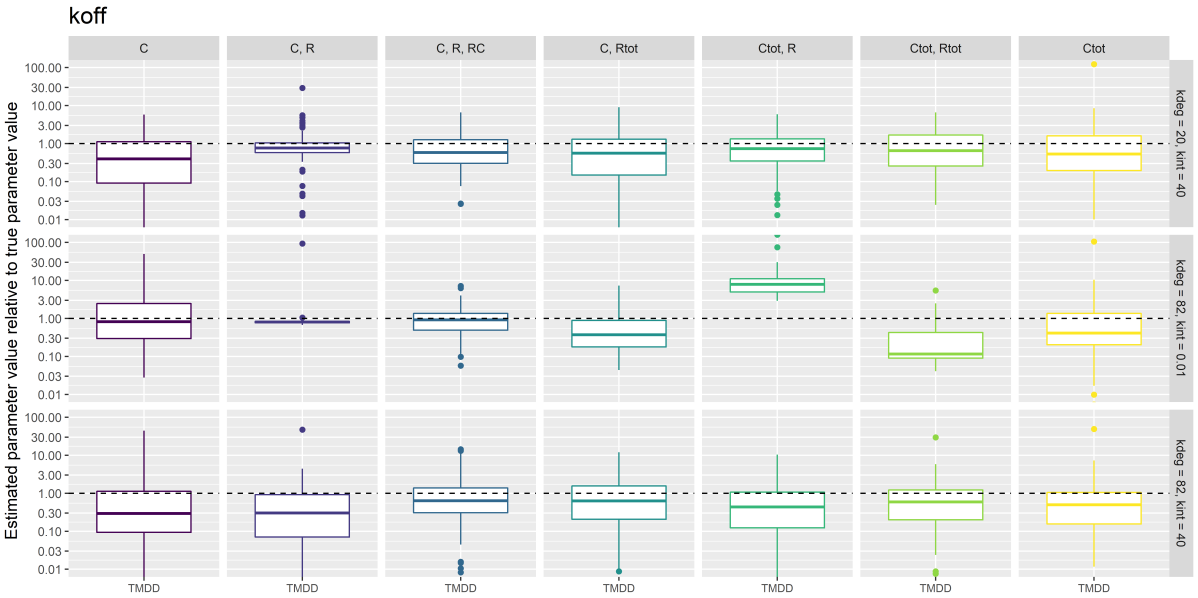

Figure 24: Plots of the estimated parameter value k_{off} (/nM/day) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

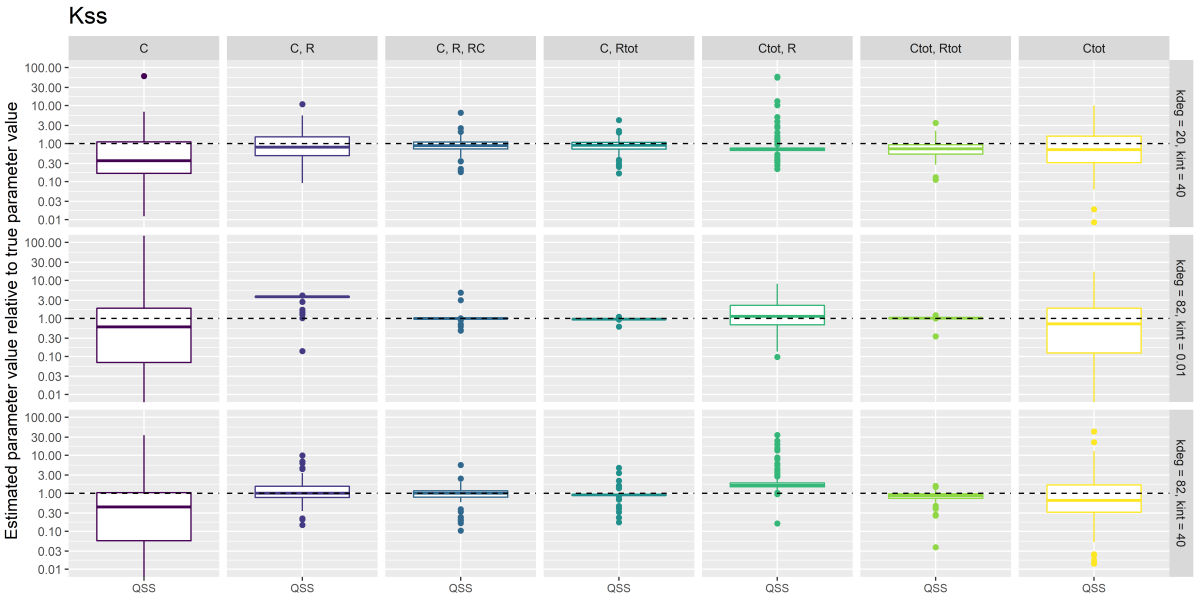

Figure 25: Plots of the estimated parameter value K_{ss} (nM) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

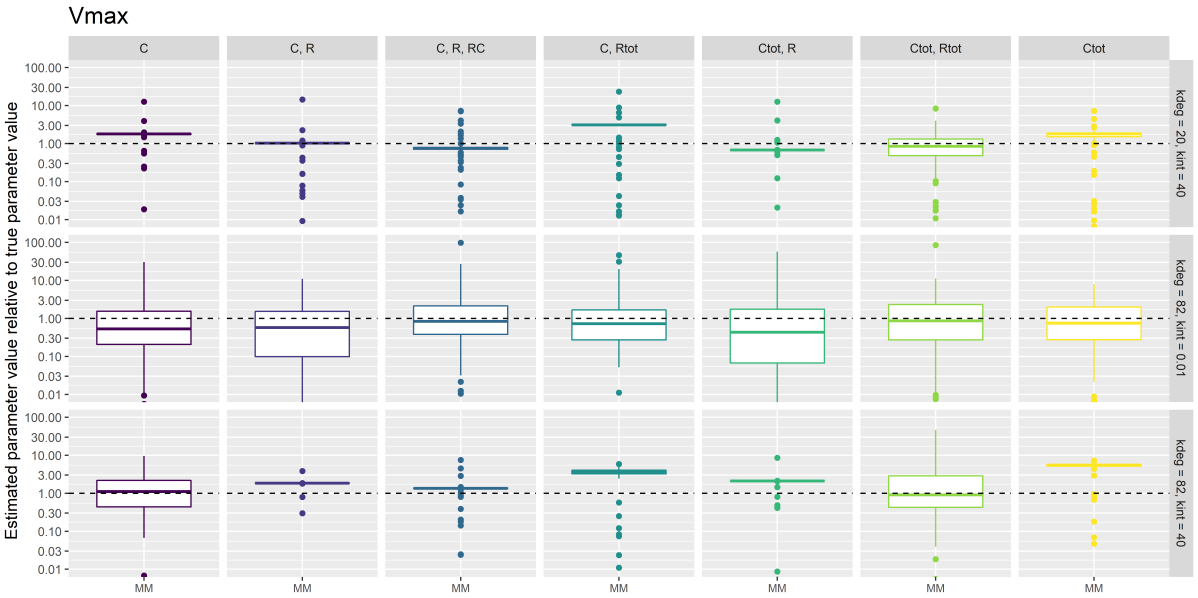

Figure 26: Plots of the estimated parameter value V_{max} (nM) relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

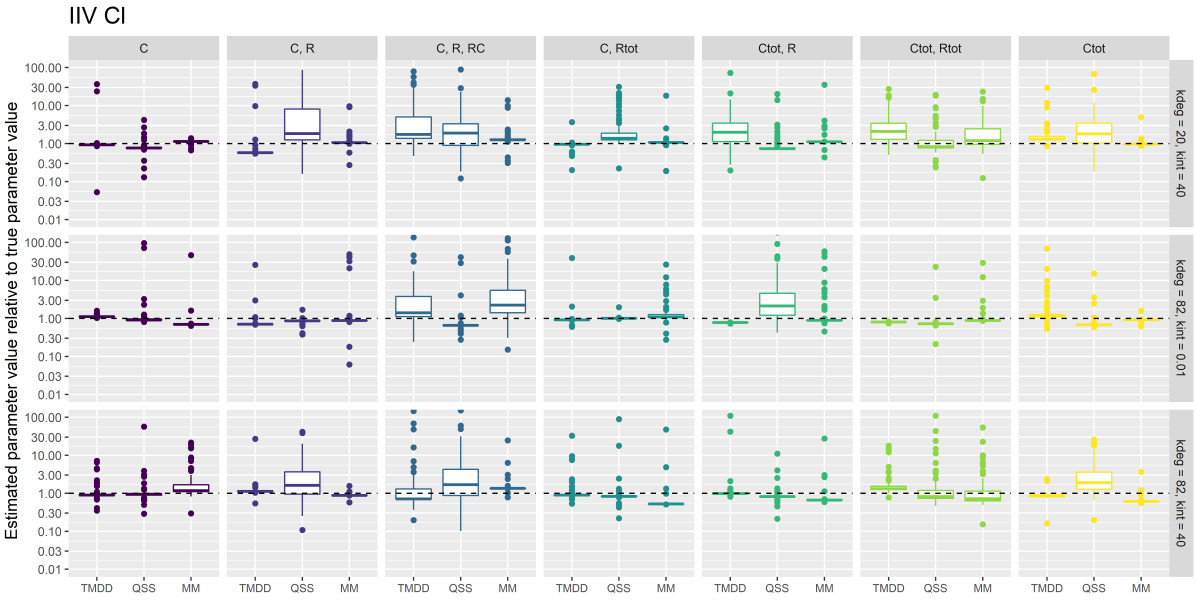

Figure 27: Plots of the estimated parameter value $IIV\ Cl$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

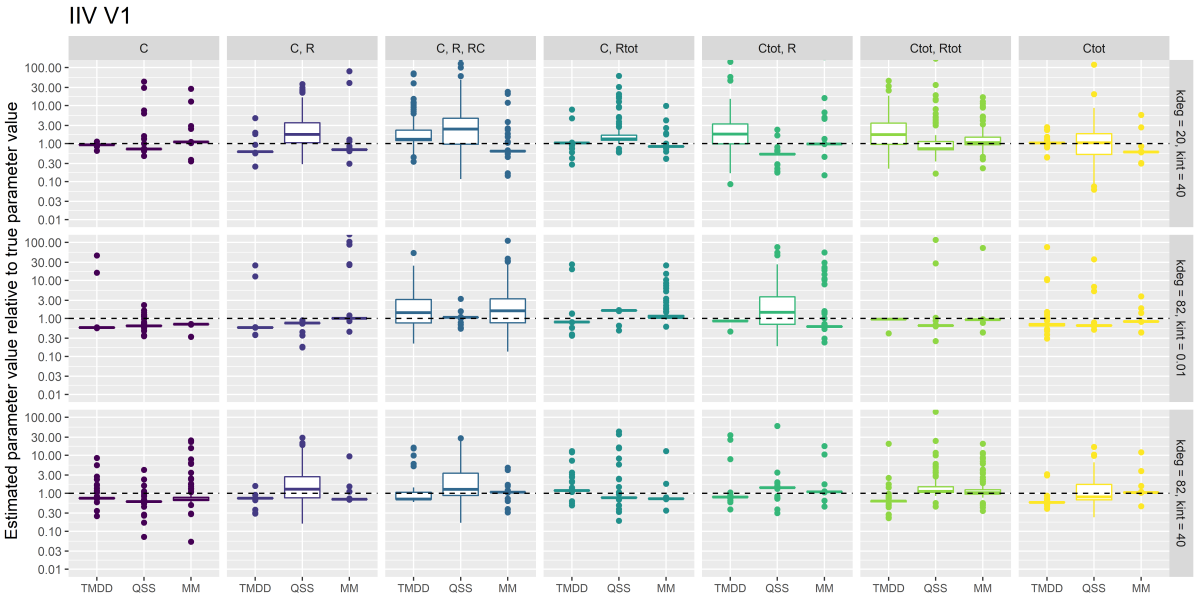

Figure 28: Plots of the estimated parameter value $IIV\ V1$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

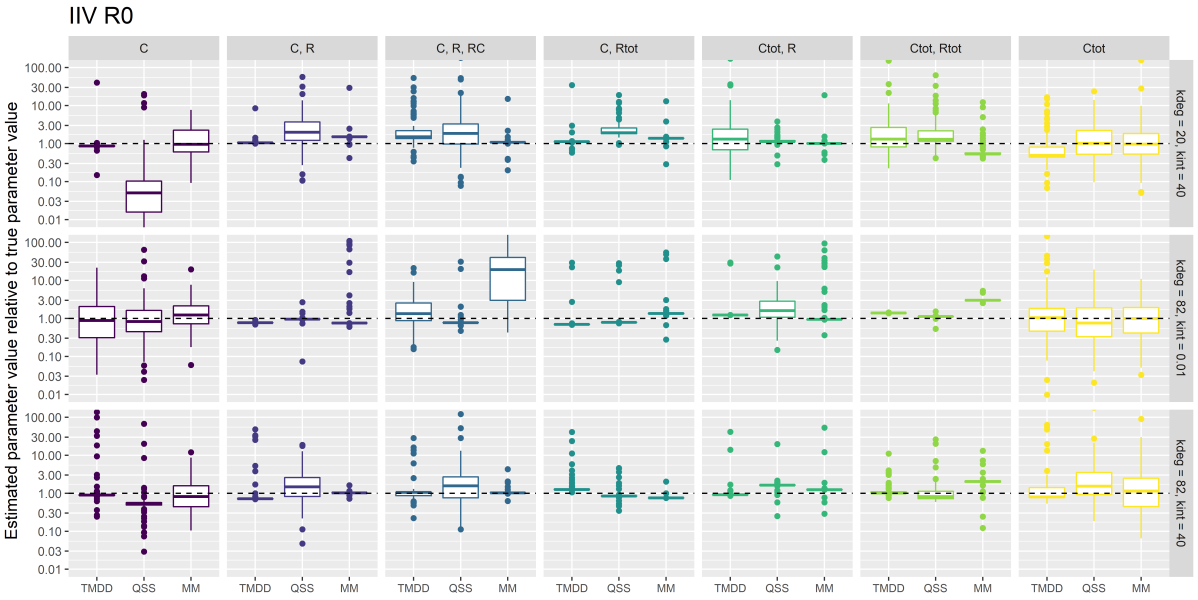

Figure 29: Plots of the estimated parameter value $IIV\ R0$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

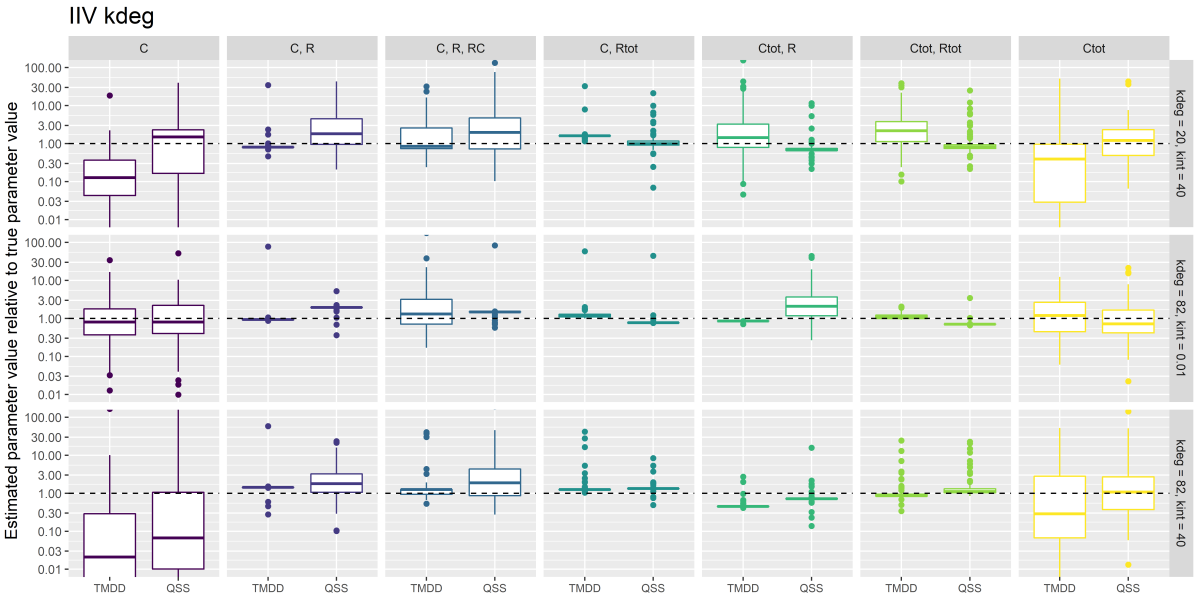

Figure 30: Plots of the estimated parameter value $IIV\ k_{deg}$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

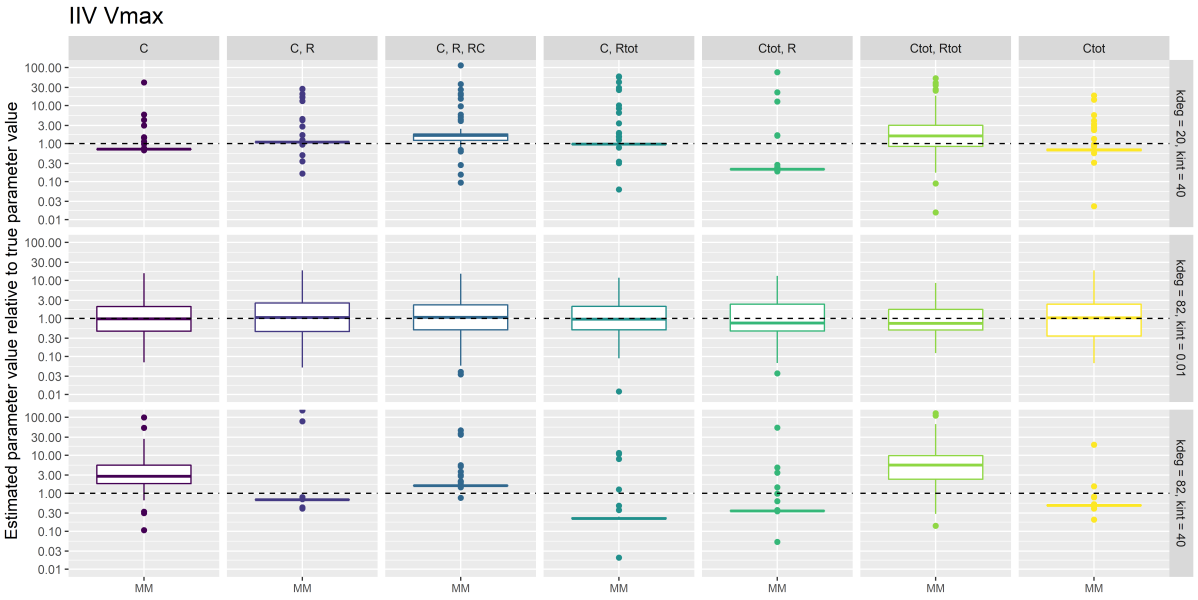

Figure 31: Plots the natural logarithm of of the estimated parameter value $IIV\ V_{max}$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

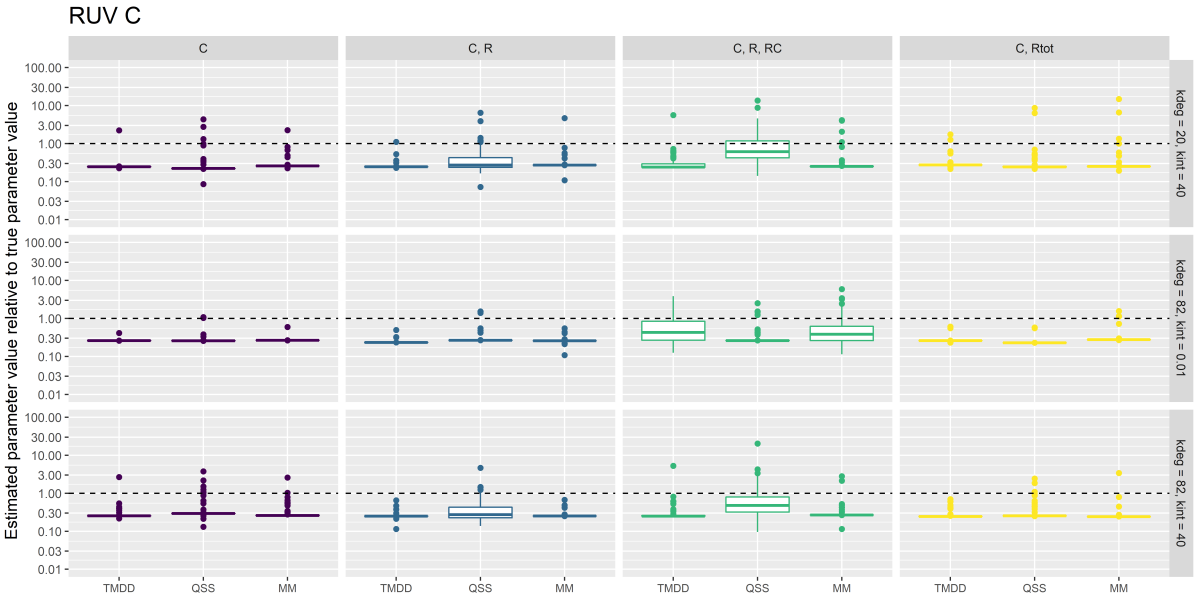

Figure 32: Plots of the estimated parameter value $RUV\ C$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

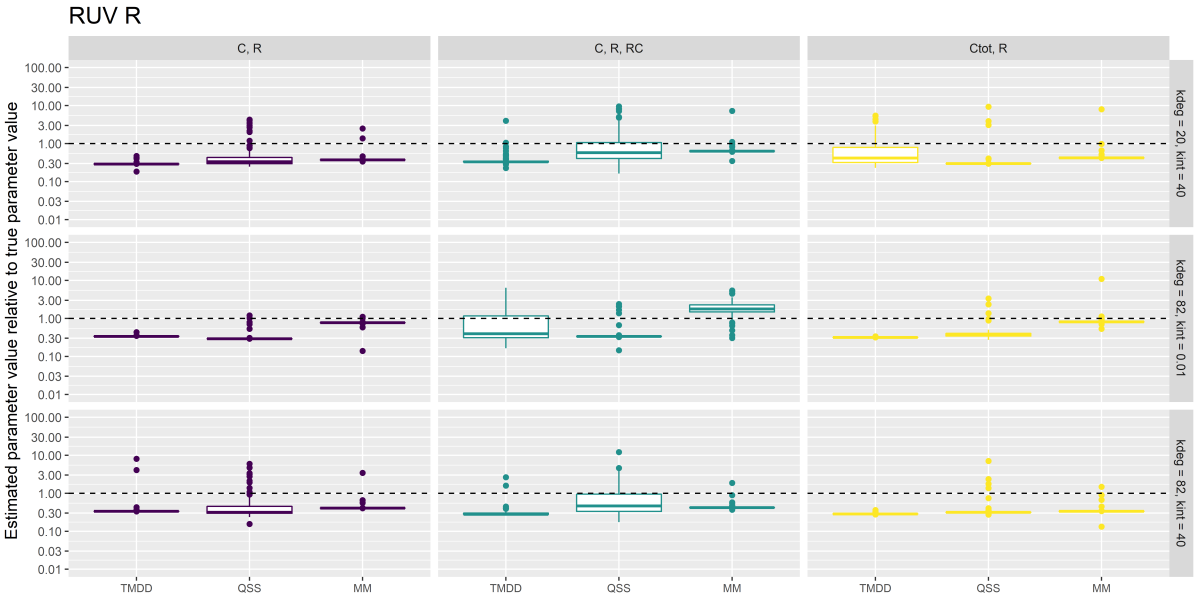

Figure 33: Plots of the estimated parameter value $RUV\ R$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

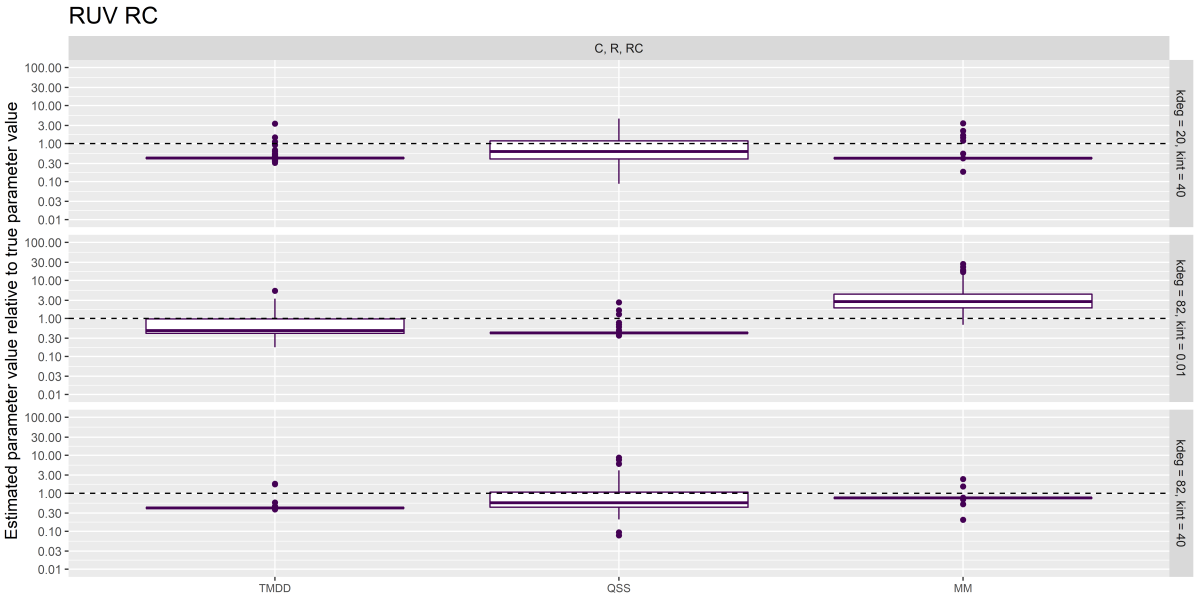

Figure 34: Plots of the estimated parameter value $RUV\ RC$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

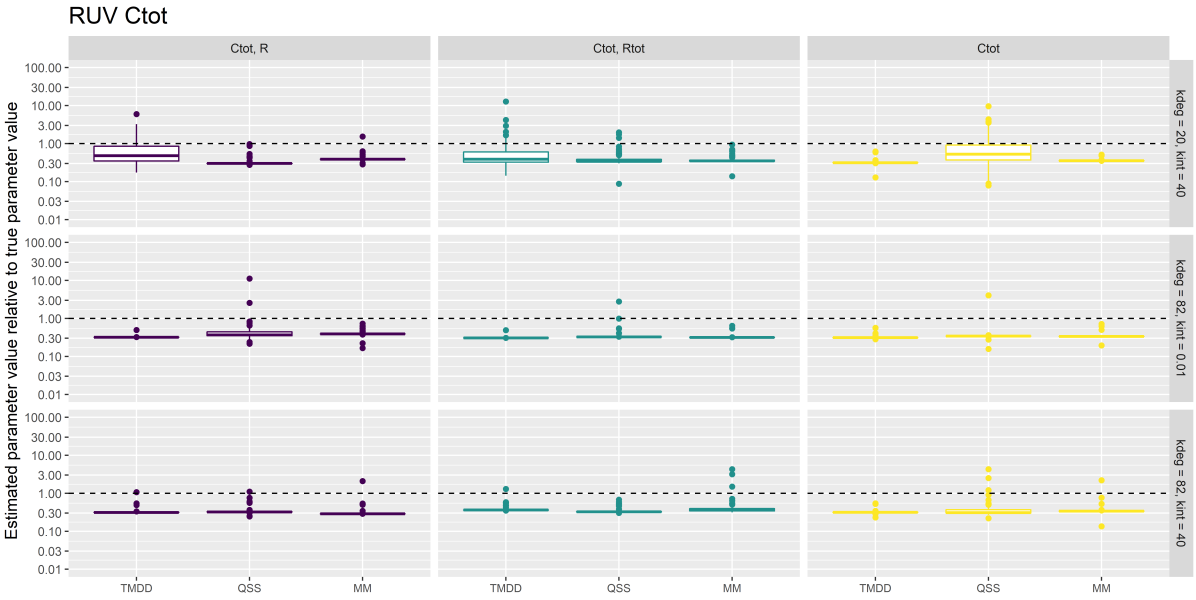

Figure 35: Plots of the estimated parameter value $RUV\ C_{tot}$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.

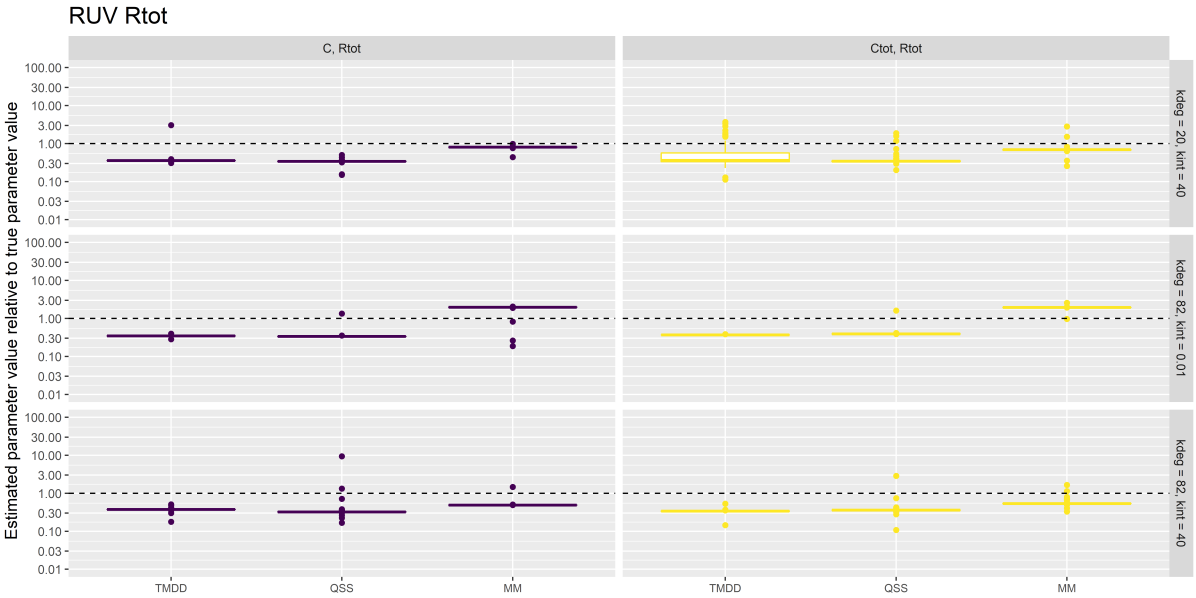

Figure 36: Plots of the estimated parameter value $RUV\ R_{tot}$ relative to the true parameter value for each model, output measure and parameter set. Boxes represent the inter-quartile range (IQR) and the solid line in the box the median value. The length of the whiskers corresponds to the furthest point from the median within 1.5 x IQR. Outliers are plotted as points. The plot y-limit has been set to [0.01,100], so not all data may be included in the plot.