ON PARAMETER ESTIMATION IN AN *IN VITRO* COMPARTMENTAL MODEL FOR DRUG-INDUCED ENZYME PRODUCTION IN PHARMACOTHERAPY

Jurjen Duintjer Tebbens, Hradec Králové, Ctirad Matonoha, Praha,
Andreas Matthios, Hradec Králové, Štěpán Papáček, České Budějovice

Received October 16, 2018. Published online March 20, 2019.

Abstract. A pharmacodynamic model introduced earlier in the literature for in silico prediction of rifampicin-induced CYP3A4 enzyme production is described and some aspects of the involved curve-fitting based parameter estimation are discussed. Validation with our own laboratory data shows that the quality of the fit is particularly sensitive with respect to an unknown parameter representing the concentration of the nuclear receptor PXR (pregnane X receptor). A detailed analysis of the influence of that parameter on the solution of the model's system of ordinary differential equations is given and it is pointed out that some ingredients of the analysis might be useful for more general pharmacodynamic models. Numerical experiments are presented to illustrate the performance of related parameter estimation procedures based on least-squares minimization.

Keywords: pharmacotherapy; pharmacodynamic modelling; constrained optimization; parameter estimation

MSC 2010: 92C45, 34A34, 65F60, 65K10

The work of Jurjen Duintjer Tebbens and Ctirad Matonoha was supported by the long-term strategic development financing of the Institute of Computer Science (RVO: 67985807) of the Czech Academy of Sciences. The work of Jurjen Duintjer Tebbens was also supported by the Czech science funding agency (GACR 17-06841S). The work of Andreas Matthios was supported by the grant agency of Charles University project GAUK 110/50/85003 and by SVV 260 414. The work of Štěpán Papáček was supported by the Ministry of Education, Youth and Sports of the Czech Republic – projects "CENAKVA" (No. CZ.1.05/2.1.00/01.0024), "CENAKVA II" (No. LO1205 under the NPU I program) and The CENAKVA Centre Development (No. CZ.1.05/2.1.00/19.0380).

1. Introduction

Efficient modelling and simulation of drug distribution profiles in organs is becoming increasingly important. This is true not only for theoretical pharmacology, where the main goal is often to describe the action of administered ligands (drugs) on the cellular level, like its influence on receptors and metabolizing enzymes, but for clinical pharmacy as well [19]. Drug concentration models help in designing patient-tailored dosing regimens and represent an important tool for the assessment of drugs safety before their approval by a state drug control institute. For instance, the US Food and Drug Administration is frequently processing computer simulation based analyses in regulatory submissions [20].

The advantages of *in silico* (numerical) experiments have for some time been routinely exploited in pharmacology, just like in other fields of health and natural sciences. Clearly, *in vitro* experiments are time consuming and demanding with respect to financial and human resources. Acquisition and preparation of chemicals, including cell cultivation, followed by drug administration and laboratory analysis to produce measurement data, is a process that can easily take several months. For *in vivo* trials the time, labor, and financial costs are even higher. *In silico* computer simulations, on the other hand, gain attractiveness as they become more reliable, affordable and user friendly [17]. Nowadays, dedicated software is available (e.g. ADAPT [1], CellDesigner [4], Simcyp [18], NONMEM [15]).

In some situations, reliable *in silico* simulations can be indispensable. There is a significant gap between the knowledge on rodent and on human drug distribution behavior because human *in vivo* experiments are often infeasible (for safety reasons, because of ethical objections, in clinical pediatrics, etc.). Often the only viable option to bridge this gap is through extrapolation of experimental data from rodents to humans using an appropriate model. Another important issue in modern pharmacology are drug-drug interactions (DDI's). DDI's are in general poorly understood and decisions are sometimes made based on a trial-and-error approach which, in the worst case, can have fatal consequences. Mathematical models that quantify drug-drug-interactions might offer useful guidance for practitioners when facing the challenges of drug administration decisions in multi-drug therapy.

The primary goal of the so-called physiologically-based pharmacokinetic (PBPK) and pharmacodynamic models is to provide time-profiles of the concentrations of the involved substances (drugs, receptors, metabolizing enzymes) in several parts of the body. This is done using compartmental models, where it is assumed that substance concentrations are distributed homogeneously over the entire compartment [2]. Examples of compartments include plasma, intracellular and extracellular fluid, adipose tissue, organs, cells, but they can represent abstract units as well [8]. The defined

compartments depend for instance on the route of administration (intravenous, oral, etc.), the target organs and the modelled ligand-induced processes. These processes are described based on the assumed bio-physical laws and take the form of ordinary differential equations (ODEs). As the involved substances usually react with several of the other substances, the result is a system of in general nonlinear differential equations whose size is at least equal to the total number of substances; substances appearing in more than one compartment have to be split accordingly and increase the size of the system. Nevertheless, the size is usually moderate (at most, say, 25) and the system can be solved numerically in reasonable time (as long as the system of ODEs is not too stiff).

A serious problem however is that not all the constants (parameters) in the system of differential equations are readily available. The model's parameters include diffusion coefficients, elimination and production rates, organ volumes, systemic clearance or blood flow rates. Some of these physiologic parameters are known from the literature or easily obtained from experimental measuring, but typically at least a small number needs to be estimated. Parameter estimation is an integral part of the PBPK and pharmacodynamic modelling process. Sometimes this is done through Monte-Carlo Markov-Chain simulations, requiring a high number of solutions of the system of ODEs [11]. Traditionally, parameter estimation is performed using the collected experimental data from donors and the subsequent curve fitting, i.e. minimization of a sum of squares based on comparing observed and model predicted concentrations. The numerical minimization procedure is in general iterative. As a consequence, in every iteration, the entire system of differential equations needs to be solved with updated values for the parameters to be estimated. Efficient numerical optimization has therefore a crucial influence on the overall computational time of the PBPK or the pharmacodynamic modelling process.

The goal of this paper is to highlight some aspects of the numerical curve fitting based parameter estimation involved in pharmacodynamic models. We will demonstrate these with a model introduced by Luke et al. [9] for xenobiotics binding to the pregnane X nuclear receptor (PXR) and inducing CYP3A4 enzymes. In the next section, this model is described in detail and its ability to predict the concentrations measured in our own laboratory experiments is discussed. In fact, the predictions are not very accurate but we discovered through trial-and-error that they can be considerably improved by doubling the value of one of the parameters as estimated in Luke et al. [9]. In Section 3 we present an analysis to explain this fact theoretically. Section 4 discusses relevance and possible consequences of the analysis for more general cases and presents related numerical experiments. The last section points out some future work and concludes the paper.

2. A compartmental model for PXR-induced CYP3A4 production

We will consider a pharmacodynamic model introduced by Luke et al. [9] for prediction of *in vitro* measurements of intracellular substance concentrations. The model for the action of a xenobiotic (here the drug rifampicin) is schematically given in Figure 1 (which also appeared in our publication [3]). We will briefly describe below the individual processes it displays.

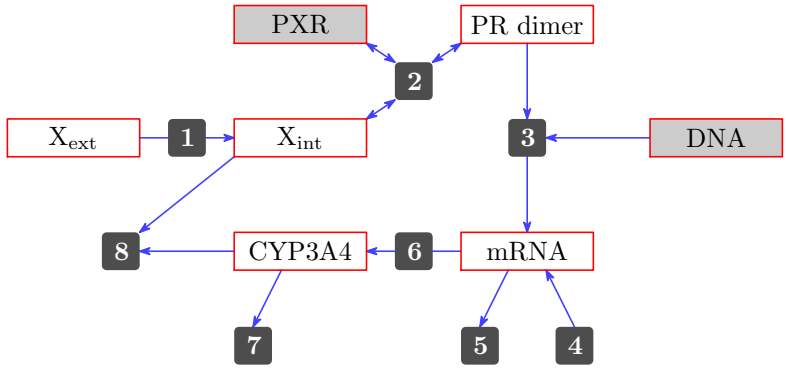

Figure 1. Schematic representation of the modelled PXR-mediated processes. Numbered squares represent the following reactions: (1) the xenobiotic enters the cell; (2) PXR binds to the xenobiotic, leading to formation of PXR/RXRα heterodimer; (3) PXR/RXRα dimer binds to DNA, increasing transcription; (4) mRNA background production; (5) degradation of mRNA; (6) the translation of mRNA forms the protein; (7) degradation of the CYP3A4 protein; (8) the CYP3A4 protein metabolizes the xenobiotic.

When the xenobiotic enters the cell,[1] it initiates several processes. A PXR/RXR (PR) heterodimer is created through binding to the nuclear receptor PXR. The heterodimer translocates to the nucleus where it stimulates, through processes described later, the production of the CYP3A4 enzyme. The enzyme, however, metabolizes not only the subtances that cause the disease, but the rifampicin as well. Besides this feedback-loop of the xenobiotic, rifampicin is also returned after dissociation of the heterodimer.

Similarly to [9], the changes of the xenobiotic concentration outside the cell $X_{ext}(t)[\mu M]$ and inside the cell $X_{int}(t)[\mu M]$ are represented by the following equations based on the assumption that the transport rate across the membrane is directly proportional to the coefficient k_{up}, to the difference in solution concentrations, and

[1] It is assumed that rifampicin enters the cell across the membrane by a simple diffusion process [21], i.e., the diffusion flux J is modelled in the compartmental framework as follows: $J = -k_{up}(X_{ext}(t) - X_{int}(t))$. The variables and parameters are described in the main text.

to the area of the membrane:

$$\frac{\mathrm{d}\,\mathrm{X}_{\mathrm{ext}}(t)}{\mathrm{d}t} = d(t) - k_{\mathrm{up}}\mathrm{X}_{\mathrm{ext}}(t) + k_{\mathrm{up}}\mathrm{X}_{\mathrm{int}}(t), \tag{2.1}$$

where the function $d(t)$[mg/min] represents the dosing added into the system, and

$$\begin{aligned}\frac{\mathrm{d}\,\mathrm{X}_{\mathrm{int}}(t)}{\mathrm{d}t} = {} & k_{\mathrm{up}}\mathrm{X}_{\mathrm{ext}}(t) - k_{\mathrm{up}}\mathrm{X}_{\mathrm{int}}(t) - k_{\mathrm{assoc}}\mathrm{X}_{\mathrm{int}}(t)(s_{\mathrm{PXR}} - \mathrm{PR}(t)) \\ & - k_{\mathrm{met}}\mathrm{CYP3A4}(t)\mathrm{X}_{\mathrm{int}}(t) + k_{\mathrm{dis}}\mathrm{PR}(t).\end{aligned} \tag{2.2}$$

Here, k_{up}[1/min] is the first order diffusion coefficient encompassing the permeability coefficient and the area of the membrane, k_{assoc}[1/min], k_{met}[1/(μMmin)], and k_{dis}[1/min] are the corresponding association, metabolization, and dissociation constants, respectively. An important parameter (as will be seen later) is the total system PXR concentration (binded and free) s_{PXR}[μM].

The change in time of the PR(t)[μM] heterodimer is described as

$$\frac{\mathrm{d}\,\mathrm{PR}(t)}{\mathrm{d}t} = k_{\mathrm{assoc}}\mathrm{X}_{\mathrm{int}}(t)(s_{\mathrm{PXR}} - \mathrm{PR}(t)) - k_{\mathrm{dis}}\mathrm{PR}(t), \tag{2.3}$$

where it is assumed that PR dissociates as well.

As the amount of rifampicin is increased, the increased concentration of the heterodimer causes it to enter the nucleus where it induces increased CYP3A4 mRNA levels. The change in time of mRNA(t)[μM] is described as

$$\frac{\mathrm{d}\,\mathrm{mRNA}(t)}{\mathrm{d}t} = k_{\mathrm{mRNA}}\mathrm{PR}(t) - k_{\mathrm{mRNA,deg}}\mathrm{mRNA}(t) + p_{\mathrm{mRNA,back}}, \tag{2.4}$$

where k_{mRNA}[1/min] is the transcription constant for mRNA. Moreover, background production and degradation of mRNA with the corresponding constants $p_{\mathrm{mRNA,back}}$[μM/min] and $k_{\mathrm{mRNA,deg}}$[1/min] are assumed.

The result of mRNA translocation into the cytoplasm is the production of the CYP3A4 enzyme. The change in time of CYP3A4(t)[μM] is described as

$$\frac{\mathrm{d}\,\mathrm{CYP3A4}(t)}{\mathrm{d}t} = k_{\mathrm{cyp}}\mathrm{mRNA}(t) - k_{\mathrm{cyp,deg}}\mathrm{CYP3A4}(t), \tag{2.5}$$

where k_{cyp}[1/min] is a translation constant for CYP3A4. Also, degradation of CYP3A4 with the corresponding constant $k_{\mathrm{cyp,deg}}$[1/min] is assumed.

Equations (2.1) through (2.5) define a system of ODEs of size five. The values of mRNA and CYP3A4 are usually given as fold induction. It means that these values

are divided by their initial (steady-state) values mRNA_{ss} and $\mathrm{CYP3A4}_{ss}$:

$$\mathrm{mRNA}_{ss} = \frac{p_{\mathrm{mRNA,back}}}{k_{\mathrm{mRNA,deg}}}, \tag{2.6}$$

$$\mathrm{CYP3A4}_{ss} = \frac{k_{\mathrm{cyp}}}{k_{\mathrm{cyp,deg}}}\mathrm{mRNA}_{ss}. \tag{2.7}$$

The initial concentrations of the remaining substances are assumed to be zero. However, in the simulations presented in [9] the dosing function $d(t)$ is not used. Instead, the administered dose is incorporated by putting the initial value of $\mathrm{X}_{\mathrm{ext}}$ equal to this dose.

Parameter	Value
k_{up}	$6.55{\cdot}10^{-3}$
k_{assoc}	$k_{\mathrm{dis}}/5.6$
k_{mRNA}	39.3
$k_{\mathrm{mRNA,deg}}$	0.04
k_{cyp}	2.5
$k_{\mathrm{cyp,deg}}$	$2.7{\cdot}10^{-4}$

Table 1. The values of known parameters.

Most of the parameters in the ODEs can be taken from previously published papers; their values are reported in Table 1. Note that the parameter k_{assoc} is computed based on the value of the unknown parameter k_{dis}. The other unknown parameters are $s_{\mathrm{PXR}}, k_{\mathrm{met}}$ and $p_{\mathrm{mRNA,back}}$. They have been estimated in [9] through curve-fitting to experimental data. Their estimated values can be found for instance in the first columns of Tables 5 and 6.

Experimental data needed for curve fitting computations often contain the concentrations of only some of the substances, like the xenobiotic concentration outside the cell $\mathrm{X}_{\mathrm{ext}}(t)$ and the CYP3A4 mRNA concentration $\mathrm{mRNA}(t)$. In particular, PXR concentrations appear to be difficult to measure.

We validated the model of Luke et al. with our own experimental data. In our experiment, primary human hepatocytes were treated with 10 μM of rifampicin. The expression level profiles of mRNA for the CYP3A4 enzyme were analyzed using the qRT-PCR method at 0, 6, 12, 24, 48, and 96 h from the beginning of the treatment. The measured average fold levels are displayed as circles in Figures 2 and 3 with the vertical lines displaying the corresponding standard deviation (every measurement was repeated twice) and the solid lines giving the levels predicted by the model of Luke et al., using our own Matlab [12] implementation. In Figure 2, we used the literature parameter values in Table 1 and for the unknown parameters we used the values as estimated by Luke et al. Clearly, our experimental data do not fit well with

the predicted concentrations. However, one of the estimated parameters, s_{PXR}, is the total amount of PXR in the cell. To our surprise, we observed by simple manual modification of this parameter, that doubling its value resulted in rather satisfactory fitting as can be seen in Figure 3 (except for the time-point 48 hours, which seems to be an outlier caused by inappropriate physical circumstances that might be ignored). This observation has also been described in [3].

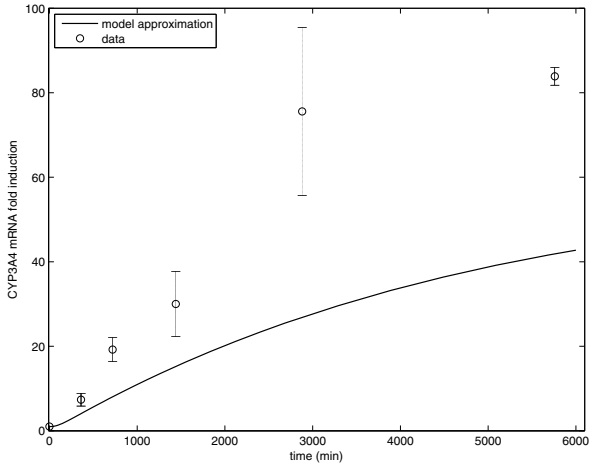

Figure 2. Time profiles of CYP3A4 mRNA induction simulated employing the model by Luke et al. [9] with an estimated total (free and bound) intracellular PXR concentration of $9.47{\cdot}10^{-7}\,\mu$M.

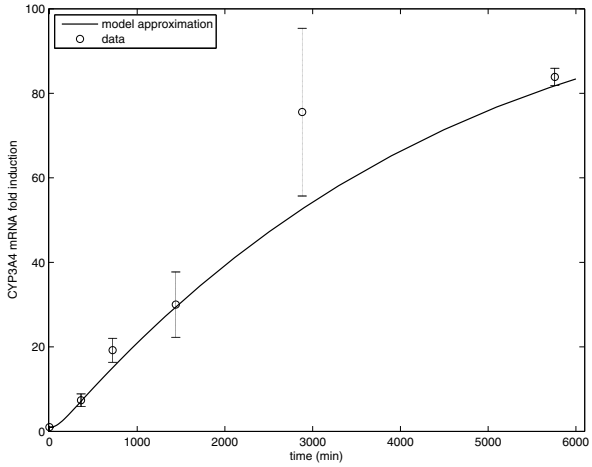

Figure 3. Time profiles of CYP3A4 mRNA induction simulated employing the model by Luke et al. [9] with an estimated total (free and bound) intracellular PXR concentration of $2 \cdot 9.47{\cdot}10^{-7}\,\mu$M.

3. Analysis

In this section we try to give an explanation for the fact that our laboratory data can be fitted rather satisfactorily by simply doubling the value for the parameter s_{PXR} as estimated in [9]. We give a straightforward analysis, which for this special situation appears to give some insight.

First we introduce new notation. The unknown substance concentrations will be collected in a size five vector x according to

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \end{pmatrix} \equiv \begin{pmatrix} \mathrm{X_{ext}}(t) \\ \mathrm{X_{int}}(t) \\ \mathrm{PR}(t) \\ \mathrm{mRNA}(t) \\ \mathrm{CYP3A4}(t) \end{pmatrix}.$$

We will also change the notation of all constants except s_{PXR} (which is the parameter of main interest) and define:

$$k_1 \equiv k_{\mathrm{up}}, \quad k_2 \equiv k_{\mathrm{assoc}}, \quad k_3 \equiv k_{\mathrm{met}}, \quad k_4 \equiv k_{\mathrm{dis}}, \quad k_5 \equiv k_{\mathrm{mRNA}},$$
$$k_6 \equiv k_{\mathrm{mRNA,deg}}, \quad k_7 \equiv p_{\mathrm{mRNA,back}}, \quad k_8 \equiv k_{\mathrm{cyp}}, \quad k_9 \equiv k_{\mathrm{cyp,deg}}.$$

I_5 will denote the identity matrix of size five, e_i its ith column. Then the system of differential equations of the previous section can be written as

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = Bx(t) + z(t), \tag{3.1}$$

with the constant matrix

$$B = \begin{pmatrix} -k_1 & k_1 & 0 & 0 & 0 \\ k_1 & -(k_1 + k_2 s_{\mathrm{PXR}}) & k_4 & 0 & 0 \\ 0 & k_2 s_{\mathrm{PXR}} & -k_4 & 0 & 0 \\ 0 & 0 & k_5 & -k_6 & 0 \\ 0 & 0 & 0 & k_8 & -k_9 \end{pmatrix} \tag{3.2}$$

representing the linear part of the system and the vector

$$z(t) = \begin{pmatrix} 0 \\ k_2 \cdot x_2(t) \cdot x_3(t) - k_3 \cdot x_2(t) \cdot x_5(t) \\ -k_2 \cdot x_2(t) \cdot x_3(t) \\ k_7 \\ 0 \end{pmatrix}$$

representing the nonlinear (quadratic) and constant parts. The initial conditions given in [9] are

$$x(0) = \begin{pmatrix} x_1(0) \\ 0 \\ 0 \\ x_4(0) \\ x_5(0) \end{pmatrix} \tag{3.3}$$

in our situation, $x_1(0)$ is the initially applied dose, i.e. $x_1(0) = 10\,\mu\text{M}$, and $x_4(0) = 7.075{\cdot}10^{-6}\,\mu\text{M}$ and $x_5(0) = 0.0655\,\mu\text{M}$ are the steady (initial) state concentrations for mRNA and CYP3A4, respectively.

In our analysis we will use the linearized version of the system of differential equations (3.1). If we define $f(x(t)) \equiv Bx(t) + z(t)$, the linearized system can be written as

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \nabla f(x(0)) \cdot x(t) + b, \quad b \equiv (\,0 \quad 0 \quad 0 \quad k_7 \quad 0\,)^{\mathrm{T}}, \tag{3.4}$$

where due to the given initial conditions $x_2(0) = x_3(0) = 0$ we have

(3.5)

$$\nabla f(x(0)) = C, \quad C \equiv \begin{pmatrix} -k_1 & k_1 & 0 & 0 & 0 \\ k_1 & -(k_1 + k_2 s_{\text{PXR}}) - k_3 x_5(0) & k_4 & 0 & 0 \\ 0 & k_2 s_{\text{PXR}} & -k_4 & 0 & 0 \\ 0 & 0 & k_5 & -k_6 & 0 \\ 0 & 0 & 0 & k_8 & -k_9 \end{pmatrix}.$$

Solutions of this linearized problem are given explicitly as

$$x_l(t) = \mathrm{e}^{Ct}(x(0) + y) - y, \quad \text{where } Cy = b. \tag{3.6}$$

For y in (3.6) we have

$$\begin{pmatrix} -k_1 & k_1 & 0 & 0 & 0 \\ k_1 & -(k_1 + k_2 s_{\text{PXR}}) - k_3 x_5(0) & k_4 & 0 & 0 \\ 0 & k_2 s_{\text{PXR}} & -k_4 & 0 & 0 \\ 0 & 0 & k_5 & -k_6 & 0 \\ 0 & 0 & 0 & k_8 & -k_9 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ k_7 \\ 0 \end{pmatrix},$$

thus

$$y = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\dfrac{k_7}{k_6} \\ \dfrac{-k_7 k_8}{k_6 k_9} \end{pmatrix}. \tag{3.7}$$

In Figures 4 and 5 we see that this solution gives satisfactory approximations of the numerical solution of the true differential equations system (3.1) for the range of interest for s_{PXR} and for the times corresponding to the first few experimental data, i.e. after 6, 12 and 24 hours. We will next concentrate on the analysis of the linearized system of differential equations (3.4) with (3.5) to explain the influence of the parameter s_{PXR} on its solution in the hope of explaining, at least for the times $t = 6 \cdot 60$, $t = 12 \cdot 60$ and $t = 24 \cdot 60$ minutes, why doubling the value estimated in [9] for s_{PXR} gives a remarkably good fit when using the true differential equations system (3.1).

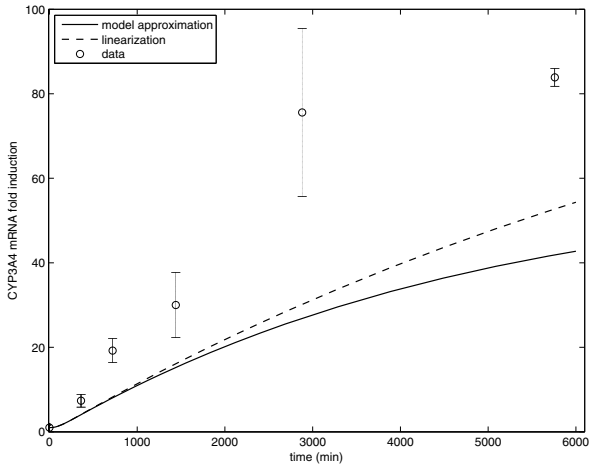

Figure 4. Time profiles of CYP3A4 mRNA induction modelled by numerical solution of (3.1) and by the solution of the linearized system (3.4) with an estimated total (free and bound) intracellular PXR concentration of $9.47 \cdot 10^{-7}$ μM.

We will use the fact that any matrix function can be expressed as a polynomial in that matrix, where the degree of the polynomial is at most the size of the matrix minus one. Hence, for the matrix exponential e^{Ct} in (3.6) we have

$$\mathrm{e}^{Ct} = p_0 I_5 + p_1 Ct + p_2 C^2 t^2 + p_3 C^3 t^3 + p_4 C^4 t^4, \tag{3.8}$$

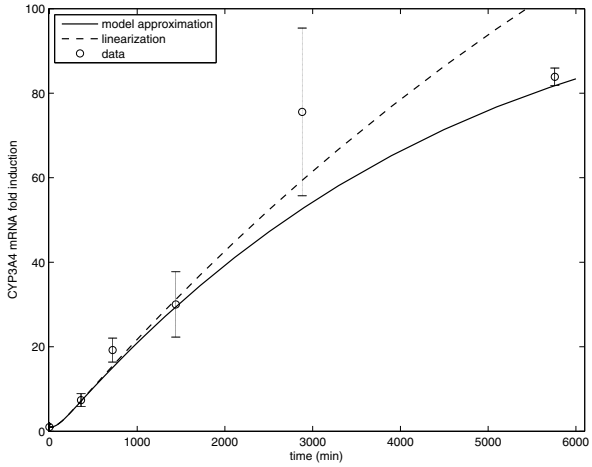

Figure 5. Time profiles of CYP3A4 mRNA induction modelled by numerical solution of (3.1) and by the solution of the linearized system (3.4) with an estimated total (free and bound) intracellular PXR concentration of $2 \cdot 9.47 \cdot 10^{-7}\,\mu$M.

see e.g. [13]. The coefficients $p_0, \ldots, p_4$ of the polynomial can be computed using for instance Sylvester's formula. They depend on t as well as on the matrix entries, including the parameter s_{PXR} we are interested in. However, the influence of s_{PXR} on these coefficients is very limited. In Table 2 we display these coefficients for e^{C_1}, where in C_1 the parameter s_{PXR} has the value $s_{\mathrm{PXR}} = 9.47 \cdot 10^{-7}\,\mu$M estimated in Luke [9], and for $\mathrm{e}^{C_{10}}$, where in C_{10} the parameter s_{PXR} has the value $s_{\mathrm{PXR}} = 9.47 \cdot 10^{-6}\,\mu$M ten times the estimate in Luke [9]. We also display the coefficients for $\mathrm{e}^{C_1 \cdot 6000}$ and $\mathrm{e}^{C_{10} \cdot 6000}$, which corresponds to the time $t = 6000$ several hours after the last laboratory observation. We see that only the last digits of the coefficients are affected if we increase the estimate for s_{PXR} with a factor 10 (when increasing with a factor two, no differences are visible in most coefficients).

	e^{C_1}	$\mathrm{e}^{C_{10}}$	$\mathrm{e}^{C_1 \cdot 6000}$	$\mathrm{e}^{C_{10} \cdot 6000}$
p_0	1.000000000000000	1.000000000000000	1.001057025103826	1.001057024996506
p_1	1.000000000000121	1.000000000000121	1.219084306280218	1.219084449821729
p_2	0.500000000731660	0.500000000731659	0.770726157661988	0.770726071609689
p_3	0.166662411269497	0.166662411269437	0.012736370268294	0.012736368725253
p_4	0.041228702593904	0.041228702592610	0.000039775431340	0.000039775426415

Table 2. Coefficients p_i in (3.8) for the original estimate of s_{PXR} (first and third column) and tenfold of the original estimate (second and fourth column), for times $t = 1$ (first two columns) and $t = 6000$ (last two columns).

In the following we will therefore neglect the influence of s_{PXR} on the p_i and restrict ourselves to the analysis of the influence of s_{PXR} on the powers of C in (3.8). For the influence of *any* paramater appearing in the matrix C in (3.5) we have the following result.

Theorem 3.1. *Neglecting the influence of parameters on the coefficients p_i in* (3.8)*, for any time t, the solution* (3.6) *of the linearized system* (3.4) *with* (3.5) *depends on any parameter k_i appearing in* (3.5)*, including s_{PXR}, according to a rational function $n(k)/d(k)$, where the numerator $n(k)$ is a polynomial of degree at most* 5 *and the denominator $d(k)$ is a polynomial of degree at most* 1.

P r o o f. Using (3.8), the term $\mathrm{e}^{Ct}x(0)$ in (3.6) equals

$$\mathrm{e}^{Ct}x(0) = (p_0 I_5 + p_1 Ct + p_2 C^2 t^2 + p_3 C^3 t^3 + p_4 C^4 t^4)x(0)$$

Clearly, every component of $\mathrm{e}^{Ct}x(0)$ is a polynomial in k_i of degree at most four. As for the second term y in (3.6), a look at the solution (3.7) shows that the constants k_6 and k_9 appear in the denominator. Hence, for these parameters we have a rational function in that parameter where the denominator is a linear function. Combination with the term $\mathrm{e}^{Ct}x(0)$ gives the claim. □

For some parameters, the influence addressed in the above theorem is described by a function simpler than a rational function. For instance, for the parameter s_{PXR} that we wish to investigate, we can decompose C in (3.5) as

$$C = \begin{pmatrix} -k_1 & k_1 & 0 & 0 & 0 \\ k_1 & -k_1 - k_3 x_5(0) & k_4 & 0 & 0 \\ 0 & 0 & -k_4 & 0 & 0 \\ 0 & 0 & k_5 & -k_6 & 0 \\ 0 & 0 & 0 & k_8 & k_9 \end{pmatrix} - \beta \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} e_2^{\mathrm{T}}, \tag{3.9}$$

where $\beta \equiv k_2 s_{\mathrm{PXR}}$ (and where we tacitly assume that k_2 is fixed). If we introduce the notation

$$C_2 \equiv \begin{pmatrix} -k_1 & k_1 & 0 & 0 & 0 \\ k_1 & -k_1 - k_3 x_5(0) & k_4 & 0 & 0 \\ 0 & 0 & -k_4 & 0 & 0 \\ 0 & 0 & k_5 & -k_6 & 0 \\ 0 & 0 & 0 & k_8 & k_9 \end{pmatrix}, \quad e_{23} \equiv \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},$$

then the term $\mathrm{e}^{Ct}x(0)$ in (3.6) can be written as

$$\begin{aligned} \mathrm{e}^{Ct}x(0) &= (p_0 I_5 + p_1 Ct + p_2 C^2 t^2 + p_3 C^3 t^3 + p_4 C^4 t^4)x(0) \\ &= p_0 x(0) + p_1 t(C_2 - \beta e_{23} e_2^{\mathrm{T}})x(0) + p_2 t^2 (C_2 - \beta e_{23} e_2^{\mathrm{T}})^2 x(0) \\ &\quad + p_3 t^3 (C_2 - \beta e_{23} e_2^{\mathrm{T}})^3 x(0) + p_4 t^4 (C_2 - \beta e_{23} e_2^{\mathrm{T}})^4 x(0), \end{aligned} \tag{3.10}$$

where we used (3.8). However, because of the special form of the initial conditions (3.3), the inner product $e_2^{\mathrm{T}}x(0)$ appearing often in (3.10) is zero. Therefore, in expressions of the form $p_i t^i (C_2 - \beta e_{23} e_2^{\mathrm{T}})^i x(0)$ in (3.10) some terms annihilate. For example,

$$\begin{aligned} p_2 t^2 (C_2 - \beta e_{23} e_2^{\mathrm{T}})^2 x(0) &= p_2 t^2 (C_2^2 - \beta C_2 e_{23} e_2^{\mathrm{T}} - \beta e_{23} e_2^{\mathrm{T}} C_2 + \beta^2 e_{23} e_2^{\mathrm{T}}) x(0) \\ &= p_2 t^2 (C_2^2 - \beta e_{23} e_2^{\mathrm{T}} C_2) x(0). \end{aligned}$$

The same can be done for all other powers of $C_2 - \beta e_{23} e_2^{\mathrm{T}}$. It is easy to see that in every case about half of the terms vanish.

In addition, for our analysis we are only interested in fitting to the experimental data measured for CYP3A4 mRNA induction, i.e. in the fourth component of the solution $x_l(t)$. This reduces the number of terms with β in (3.10) further. For example,

$$e_4^{\mathrm{T}} p_2 t^2 (C_2 - \beta e_{23} e_2^{\mathrm{T}})^2 x(0) = p_2 t^2 e_4^{\mathrm{T}} (C_2^2 - \beta e_{23} e_2^{\mathrm{T}} C_2) x(0) = p_2 t^2 (e_4^{\mathrm{T}} C_2^2 x(0)).$$

Doing this for all other powers of $C_2 - \beta e_{23} e_2^{\mathrm{T}}$ as well, we obtain the following simplification.

Proposition 3.1. *Neglecting the influence of β on the coefficients p_i in (3.8), the fourth component of the vector $\mathrm{e}^{Ct}x(0)$ in (3.6) depends on β quadratically as follows:*

$$\begin{aligned} e_4^{\mathrm{T}} \mathrm{e}^{Ct} x(0) &= e_4^{\mathrm{T}} (p_0 I_5 + p_1 C t + p_2 C^2 t^2 + p_3 C^3 t^3 + p_4 C^4 t^4) x(0) \\ &= p_0 e_4^{\mathrm{T}} x(0) + p_1 t (e_4^{\mathrm{T}} C_2 x(0)) + p_2 t^2 (e_4^{\mathrm{T}} C_2^2 x(0)) + p_3 t^3 (e_4^{\mathrm{T}} C_2^3 x(0)) \\ &\quad - p_3 t^3 \beta (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) + p_4 t^4 (e_4^{\mathrm{T}} C_2^4 x(0)) \\ &\quad - \beta p_4 t^4 (e_4^{\mathrm{T}} C_2^2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) - \beta p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2^2 x(0)) \\ &\quad + \beta^2 p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)). \end{aligned}$$

P r o o f. The result follows by writing out all expressions of the form $p_i t^i (C_2 - \beta e_{23} e_2^{\mathrm{T}})^i x(0)$ for $i = 0, \ldots, 4$ and exploiting the fact that $e_2^{\mathrm{T}} x(0) = 0$ and $e_4^{\mathrm{T}} e_{23} = 0$. □

Analogous simplifications can be carried out for the influence of some other parameters, like, for example, k_4.

The second term y in (3.6) does not depend on β at all, see (3.7). Summarizing, for the fourth component of the solution of the linearized model, which approximates

the true model simulation for CYP3A4 mRNA induction, we obtain the following corollary.

Corollary 3.1. *Neglecting the influence of β on the coefficients p_i in* (3.8)*, the fourth component of the solution $e_4^{\mathrm{T}} x_l(t)$ of the linearized system* (3.4) *with* (3.5) *depends quadratically on β as follows:*

$$\begin{aligned} e_4^{\mathrm{T}} x_l(t) = {} & \beta^2 p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) - \beta p_3 t^3 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) \\ & - \beta p_4 t^4 (e_4^{\mathrm{T}} C_2^2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) - \beta p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2^2 x(0)) \\ & - y_4 + p_0 e_4^{\mathrm{T}} x(0) + p_1 t (e_4^{\mathrm{T}} C_2 x(0)) + p_2 t^2 (e_4^{\mathrm{T}} C_2^2 x(0)) \\ & + p_3 t^3 (e_4^{\mathrm{T}} C_2^3 x(0)) + p_4 t^4 (e_4^{\mathrm{T}} C_2^4 x(0)). \end{aligned}$$

P r o o f. The claim follows from $e_4^{\mathrm{T}} x_l(t) = e_4^{\mathrm{T}} \mathrm{e}^{Ct} x(0) - e_4^{\mathrm{T}} y$ and from Proposition 3.1. □

In the first two rows of Table 3 the observed, laboratory CYP3A4 mRNA average fold induction values for the times $t_1 = 6{\cdot}60$, $t_2 = 12{\cdot}60$, $t_3 = 24{\cdot}60$, $t_4 = 48{\cdot}60$ and $t_5 = 96{\cdot}60$ minutes are displayed with their standard deviations. Let us denote them by $O(t_i)$ for the given time points. As mentioned, fold induction values are relative values with respect to steady state. In the model of Luke et al. they are predicted by the values $\mathrm{mRNA}(t_i)/\mathrm{mRNA}_{SS}$, with the steady-state concentration $\mathrm{mRNA}_{SS} = 7.075{\cdot}10^{-6}\,\mu\mathrm{M}$. For particular estimates of the parameter s_{PXR} and for the given time points, these predictions are displayed in the remaining rows of Table 3.

time (min)	$t_1 = 360$	$t_2 = 720$	$t_3 = 1440$	$t_4 = 2880$	$t_5 = 5760$
$O(t)$	7.36	19.23	30.01	75.57	83.88
stand. dev. for $O(t)$	± 1.51	± 2.84	± 7.72	± 19.86	± 2.09
$s_{\mathrm{PXR}} = 9.47{\cdot}10^{-7}$	4.06	8.04	15.24	26.87	41.87
$s_{\mathrm{PXR}} = 2{\cdot}9.47{\cdot}10^{-7}$	7.12	15.08	29.46	52.63	81.79
s_{PXR}: solution of (3.11)	7.285	18.724	28.322	66.895	66.027
s_{PXR}: solution of (3.12)	7.069	14.974	29.239	52.231	81.172

Table 3. CYP3A4 mRNA induction: Observed values (first row), their standard deviations (second row) and values predicted (remaining rows) with the model (3.1) for several choices of the parameter s_{PXR}.

Using our linearized model, the observed fold induction values can be predicted as

$$\frac{e_4^{\mathrm{T}} x_l(t_i)}{\mathrm{mRNA}_{SS}},$$

where $x_l(t)$ is given in (3.6). If we wish to find the parameter β for which our *linearized* model fits the observed data best at a given time point, we can solve the quadratric equation

$$q_2(t_i)\beta^2 + q_1(t_i)\beta + q_0(t_i) = \mathrm{mRNA}_{SS}O(t_i), \tag{3.11}$$

with coefficients

$$\begin{aligned}
q_0(t) &= y_4 + p_0 e_4^{\mathrm{T}} x(0) + p_1 t(e_4^{\mathrm{T}} C_2 x(0)) + p_2 t^2 (e_4^{\mathrm{T}} C_2^2 x(0)) \\
&\quad + p_3 t^3 (e_4^{\mathrm{T}} C_2^3 x(0)) + p_4 t^4 (e_4^{\mathrm{T}} C_2^4 x(0)) \\
q_1(t) &= -p_3 t^3 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) - p_4 t^4 (e_4^{\mathrm{T}} C_2^2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)) \\
&\quad - p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2^2 x(0)) \\
q_2(t) &= p_4 t^4 (e_4^{\mathrm{T}} C_2 e_{23})(e_2^{\mathrm{T}} C_2 x(0)),
\end{aligned}$$

see Corollary 3.1. We note that the coefficients p_i depend on t; we have omitted this dependence for simplicity of presentation.

Table 4 shows the solutions of the quadratic equation in (3.11) for the time points used in the *in vitro* experiment. It displays the corresponding values for s_{PXR} as well. We see that these values are close (in particular for the time points 6 and 24 hours) to the double value $s_{\mathrm{PXR}} = 2{\cdot}9.47{\cdot}10^{-7} = 1.894{\cdot}10^{-6}$ of the estimate obtained in [9], which gave a surprisingly good fit, see Figure 3. The predictions that the individual solutions of (3.11) give with the original, nonlinearized model (3.1) are displayed in the last but one row of Table 3. They yield lower predictions than the linearized model, which is what one would expect seeing the curves in Figures 2 and 3. For most time points they fall inside the corresponding observation plus or minus the standard deviation. But as only a single value for s_{PXR} can be used for all time-points, the quality of these predictions is of limited importance.

time (min)	$t_1 = 360$	$t_2 = 720$	$t_3 = 1440$	$t_4 = 2880$	$t_5 = 5760$
β	$3.58{\cdot}10^{-11}$	$4.38{\cdot}10^{-11}$	$3.34{\cdot}10^{-11}$	$4.45{\cdot}10^{-11}$	$2.79{\cdot}10^{-11}$
s_{PXR}	$1.95{\cdot}10^{-6}$	$2.38{\cdot}10^{-6}$	$1.82{\cdot}10^{-6}$	$2.42{\cdot}10^{-6}$	$1.52{\cdot}10^{-6}$

Table 4. Optimal values of β and s_{PXR} from the solution of (3.11) for individual time points.

If we wish to find an appropriate value of β (and hence s_{PXR}) over all observed times, we can solve the minimization problem

$$\min_{\beta} F(\beta), \quad F(\beta) = \left\| \begin{pmatrix} q_2(t_1)\beta^2 + q_1(t_1)\beta + q_0(t_1) - \mathrm{mRNA}_{SS}O(t_1) \\ q_2(t_2)\beta^2 + q_1(t_2)\beta + q_0(t_2) - \mathrm{mRNA}_{SS}O(t_2) \\ q_2(t_3)\beta^2 + q_1(t_3)\beta + q_0(t_3) - \mathrm{mRNA}_{SS}O(t_3) \\ q_2(t_4)\beta^2 + q_1(t_4)\beta + q_0(t_4) - \mathrm{mRNA}_{SS}O(t_4) \\ q_2(t_5)\beta^2 + q_1(t_5)\beta + q_0(t_5) - \mathrm{mRNA}_{SS}O(t_5) \end{pmatrix} \right\|. \tag{3.12}$$

We can solve this equation in a least-squares manner. The solution then is $\beta = 3.456{\cdot}10^{-11}$, hence $s_{\mathrm{PXR}} = \beta/k_2 = 1.879{\cdot}10^{-6}$. This is even closer to the double value $s_{\mathrm{PXR}} = 2{\cdot}9.47{\cdot}10^{-7} = 1.894{\cdot}10^{-6}$ of the estimate obtained in [9] than any of the individual solutions of (3.11) and thus offers an explanation for the fact that the double value $s_{\mathrm{PXR}} = 2{\cdot}9.47{\cdot}10^{-7}$ gives rather accurate predictions. The last row of Table 3 displays the predictions obtained with the original, nonlinearized model (3.1), where $s_{\mathrm{PXR}} = 1.879{\cdot}10^{-6}$. They are of course very close to the predictions with the double value $s_{\mathrm{PXR}} = 2{\cdot}9.47{\cdot}10^{-7} = 1.894{\cdot}10^{-6}$. In fact they are slightly smaller, which can be explained by the fact that we found an appropriate value for the linearized model, which tends to give higher predictions than the original model.

4. Discussion and complementing numerical experiments

The analysis of the previous section is clearly tailored to a specific situation for answering a particular question, which concerns merely one parameter to be estimated. Nevertheless, some more general observations for curve-fitting based parameter estimation in pharmacodynamic models can be derived from it. Below we list some of them.

- ▷ Assuming a linearized version of the given pharmacodynamic model (i.e. an ODE of the form (3.4)) yields good approximations of the predictions of the model itself, we can analyze a system of linear ODEs where the system matrix $C = \nabla f(x(0))$ is usually small, because only a limited number of substances is typically involved (at most, say, 20–25).
- ▷ The entries of C often depend linearly on the parameters of the model. This is true, because many pharmacologic processes are described as zero or first-order reactions, such as simple diffusion, membrane transport or degradation. An exception is given by processes modelled using Michaelis-Menten kinetics (or a generalization, the Hill-Langmuir equation), where a parameter appears in the denominator.
- ▷ With a small matrix size and linear dependence on parameters, the entries of the term $\mathrm{e}^{Ct}x(0)$ in the solutions (3.6) of the linearized model will approximately be polynomials of low degree in the parameters. This follows from the fact that any size n matrix function is a polynomial in that matrix of degree at most $n-1$ (a consequence of the Cayley-Hamilton theorem). However, here we assume that the dependence of the coefficients of the polynomial on the parameters is negligible. The entries of the term y of (3.6), however, can be rational functions of some parameters.
- ▷ In typical situations, only a few of the parameters need to be estimated. Moreover, these will be very sparsely spread over the matrix as they are involved in one of

the modelled pharmacologic (or biophysical) processes only. They can appear on several rows if several substances are involved, but in general, the unknown parameters can be isolated in a small rank-modification of the matrix, as was done in (3.9).

- ▷ Several of the initial substance concentrations will be zero, i.e. several components of $x(0)$ are zero. In combination with the isolation from C of unknown parameters in a small-rank matrix, the vector $\mathrm{e}^{Ct}x(0)$ may depend on unknown parameters according to a polynomial of degree significantly lower than $n-1$ (neglecting the influence of that parameter on the polynomial representing e^{Ct}).
- ▷ Data fitting is done with respect to the observed concentrations for some substances only (other substances may be very difficult to measure *in vitro*). A further reduction of the degree of polynomial dependence holds for unknown parameters appearing in the rows of C *not* corresponding to the row of the substance used for data fitting. In favorable cases, a near-optimal estimate for some parameters with respect to the linearized model can thus be found analytically. This may give useful initial guesses for the original pharmacodynamic model.
- ▷ The assumption that a linearized version of the pharmacodynamic model yields good approximations of the predictions of the model itself can be too strong or may hold only for the initial say 24 hours of the experiment. We remark, however, that the numerical solution of the model, yielding its predictions, is sometimes constructed through subsequent linearization after each time point.
- ▷ The parameter estimation problem considered here is an idealized problem in the sense that additional physical restrictions should be incorporated. Most parameters are not allowed to be negative and are physically meaningful in a particular interval only. The corresponding optimization problem represents therefore in fact constraint optimization. Also, fitting to *average* observed concentrations (the circles in Figures 2 and 3) does not take into account the natural deviation in repeated laboratory experiments (measurement error); fitting requirements could be relaxed. Finally, we modelled here with the actual physical parameter values; scaling them to avoid differences of many orders of magnitude between some parameters might be appropriate.

In the remainder of this section we perform some complementary experiments to further assess the dependence of the predicted CYP3A4 mRNA fold induction $x_4(t)$ on the parameters $s_{\mathrm{PXR}}, k_1, \ldots, k_9$. In contrast with the previous section we focus on iterative numerical optimization of all unknown parameters simultaneously. Instead of computations in Matlab [12], in this section most computations were done in Fortran [5], but some, for comparison, were done in ADAPT [1], which is an example of pharmakokinetic/dynamic systems software popular among clinical

pharmacists. We recall that the experimental data are reported in the first row of Table 3. The value at the time $t_4 = 48{\cdot}60 = 2880$ is somewhat outlying and might be ignored. This is another typical aspect of parameter estimation in pharmacodynamic modelling. In some cases the use of robust statistics for outlier detection may be very beneficial.

First, we describe our Fortran experiments. An overview of the known parameters involved in our model was given in Table 1. The parameters to be estimated are s_{PXR}, k_3, k_4, k_7. For this purpose we introduce the vector $k = (s_{\mathrm{PXR}}, k_3, k_4, k_7)^{\mathrm{T}} \in \mathbb{R}^4$ and define the ordinary least squares (OLS) function

$$J(k) = \sum_{i=0}^{5} (x_4(t_i, k) - O(t_i))^2, \tag{4.1}$$

where $x(t)$ is the solution of (3.1); note that due to the dependence of $x_4(t)$ on the parameters k we use the notation $x_4(t, k)$. For solving the involved system of ODEs we used the ODEPACK solver [7], written in Fortran, with the option $MF = 21$. In this case, an implicit backward differentiation formula method and a modified Newton iteration with user-supplied analytical Jacobian are used, see [7] for more details. We remark that the choice of parameters influences the stiffness of the ODE as they are the ratio of smallest to largest eigenvalue real part of the matrix of the linearized problem.

We will compare the quality of several parameter estimate choices using the achieved value of the OLS function J. One choice are the values obtained in [9] (based on minimization of an OLS function with different observed data), this is approx # 1 in Table 5. Another choice, discussed in detail in the previous section, differs in that we double the value of the parameter s_{PXR} (approx # 2 in Table 5). Next, we minimize the function J over all parameters in k with an optimization algorithm (approx # 3 in Table 5), hence we numerically solve the problem

$$\min_{k \in \mathbb{R}^4} J(k).$$

Finally, as the value $O_4(t_4)$ looks as an outlier, we performed minimization of J without this point, i.e., the fourth summand in (4.1) is considered zero (approx # 4 in Table 5).

For minimization of the function J, the UFO optimization software library [10], written in Fortran, was used. Starting from an initial guess $k^{(0)}$, a sequence of iterations $k^{(l+1)} = k^{(l)} + \alpha^{(l)} d^{(l)}$ is constructed. Here, $d^{(l)}$ is a direction vector determined on the basis of the values $d^{(\lambda)}$, $J(k^{(\lambda)})$, $\nabla J(k^{(\lambda)})$ for $\lambda = 1, \ldots, l-1$, and $\alpha^{(l)}$ is a steplength (automatically) determined on the basis of the behavior of J in a neighborhood of the point $k^{(l)}$. Note that the gradients $\nabla J(k^{(\lambda)})$ are computed

numerically using finite differences. In all cases, we used Luke's parameter values as the starting point $k^{(0)}$ (displayed in the first column of Table 5).

In addition to numerical experiments where, as in [9], no dosing function is used, i.e. $d(t) = 0$ with the initial value $x_1(0) = 10\,\mu$M, we also performed numerical experiments, where the dosing of $10\,\mu$M during the first hour is translated by putting $d(t) = 1/6\,\mu\text{M}/\text{ min}$ if $t < 60$ and initial value $x_1(0) = 0$. The results of this second type of computations are displayed in Table 6 (with approx # $5, 6, 7$ and 8 corresponding, respectively, to approx # $1, 2, 3$ and 4 in Table 5). We note that having $x_1(0) = 0$ can further simplify the analysis like the one given in the previous section. Tables 5–6 summarize the final parameters used to generate the approximation of the CYP3A4 mRNA fold induction, i.e. the function $x_4(t, k)$ together with the value of the OLS function J.

approx #	1	2	3	4
parameters	Luke [9]	double s_{PXR} [3]	(4.1) minimal	without outlier
s_{PXR}	$9.47{\cdot}10^{-7}$	$18.94{\cdot}10^{-7}$	$9.22{\cdot}10^{-7}$	$9.31{\cdot}10^{-7}$
k_3	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$
k_4	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$
k_7	$2.83{\cdot}10^{-7}$	$2.83{\cdot}10^{-7}$	$1.22{\cdot}10^{-7}$	$1.36{\cdot}10^{-7}$
$J(k)$	4491.11	547.96	366.78	15.29

Table 5. Parameter values for $d(t) = 0$ and $x_1(0) = 10\,\mu$M.

approx #	5	6	7	8
parameters	Luke [9]	double s_{PXR} [3]	(4.1) minimal	without outlier
s_{PXR}	$9.47{\cdot}10^{-7}$	$18.94{\cdot}10^{-7}$	$9.36{\cdot}10^{-7}$	$9.31{\cdot}10^{-7}$
k_3	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$	$2.47{\cdot}10^{-5}$
k_4	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$	$1.03{\cdot}10^{-4}$
k_7	$2.83{\cdot}10^{-7}$	$2.83{\cdot}10^{-7}$	$1.25{\cdot}10^{-7}$	$1.36{\cdot}10^{-7}$
$J(k)$	4539.04	575.59	378.80	21.20

Table 6. Parameter values for $d(t) = 1/6$ and $x_1(0) = 0\,\mu$M.

From the last row of both tables we can see that doubling the parameter s_{PXR} significantly improves the results. Not surprisingly, however, the best results (the smallest value of the function J) are obtained using minimization of J and optimizing all parameters s_{PXR}, k_3, k_4, k_7 at the same time. The very low values of $J(k)$ in the last columns are explained by the fact that the contribution of the time point $t_4 = 48{\cdot}60$, which as an apparent outlier is the largest, is missing in the sum of squares (4.1). Figures 6–7 show the corresponding curves of $x_4(t, k)$. Whereas the curves for approx # 3 and 7 show the best global fit, the best fit with the apparent

outlier omitted is for the curves for approx # 4 and 8. Nevertheless, the curves for approx # 2 and 6 come remarkably close. Hence doubling the parameter s_{PXR} seems to give an estimate that is robust with respect to outliers. One might even be tempted to think this suggests that the doubled parameter value $s_{\mathrm{PXR}} = 18.94{\cdot}10^{-7}$ is close to the actual physically valid value. But such a claim is hard to make: The interdonor variability of total PXR concentration (i.e. of s_{PXR}) may be large; after all, the originally estimated value $s_{\mathrm{PXR}} = 9.47{\cdot}10^{-7}$ yielded a satisfactory fit for other observed data, namely those that were used for curve-fitting in [9].

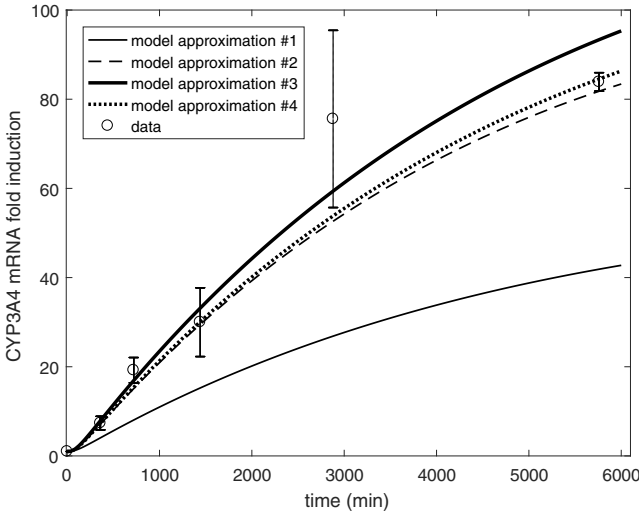

Figure 6. CYP3A4 mRNA fold induction $x_4(t,k)$ for $d(t) = 0$ and $x_1(0) = 10\,\mu$M.

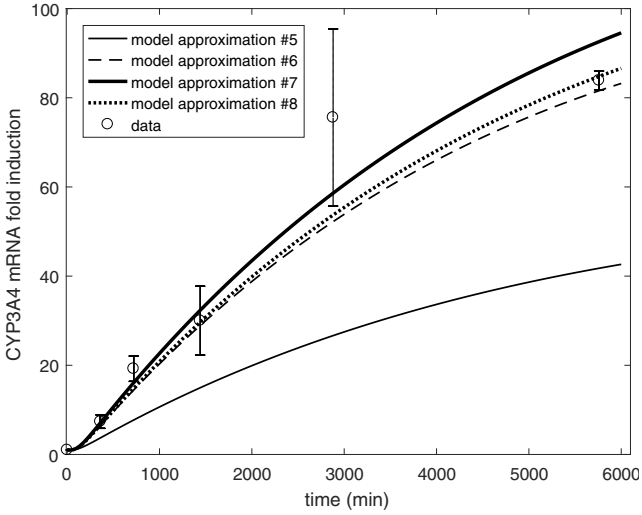

Figure 7. CYP3A4 mRNA fold induction $x_4(t,k)$ for $d(t) = 1/6$ and $x_1(0) = 0\,\mu$M.

From the results in Tables 5–6 we can deduce several other conclusions. First, it seems that only the parameters s_{PXR} and k_7 influence the results qualitatively (the parameters k_3 and k_4 are nearly identical over all columns). Concerning the situation when there is no dosing, $d(t) = 0$, and initial $x_1(0) = 10\,\mu$M and the situation when there is administration of $10\,\mu$M during the first hour (i.e. $d(t) = 1/6\,\mu\mathrm{M}/\,\mathrm{min}$) and initial $x_1(0) = 0$, one can see no significant difference in the optimal parameters and curves. However, we report that the number of function evaluations (i.e. of ODE runs) performed during the numerical iterative optimization processes were 168 and 148 for approximations # 3 and 4, respectively, whereas these were 2174 and 2202 for approximations # 7 and 8, respectively. This can be explained by the discontinuity of $d(t)$ for approximations # 7 and 8: At the moment when t equals 60 minutes it jumps from $1/6$ to 0.

It may seem contradictory that the curves for approximations #2 and #4 (and for # 6 and # 8) are very close while the corresponding parameters s_{PXR} and k_7 are rather different. To understand this phenomenon, we performed two additional numerical tests consisting in producing curves with the value for s_{PXR} fixed while changing the value of k_7 and vice versa. The results are depicted in Figures 8–9. One can deduce that for fixed s_{PXR}, the CYP3A4 mRNA fold induction curve is higher with increasing k_7, while for fixed k_7, this curve is higher with decreasing s_{PXR}. Therefore, a curve with twice a greater value of s_{PXR} can be close to one with twice a smaller value of k_7.

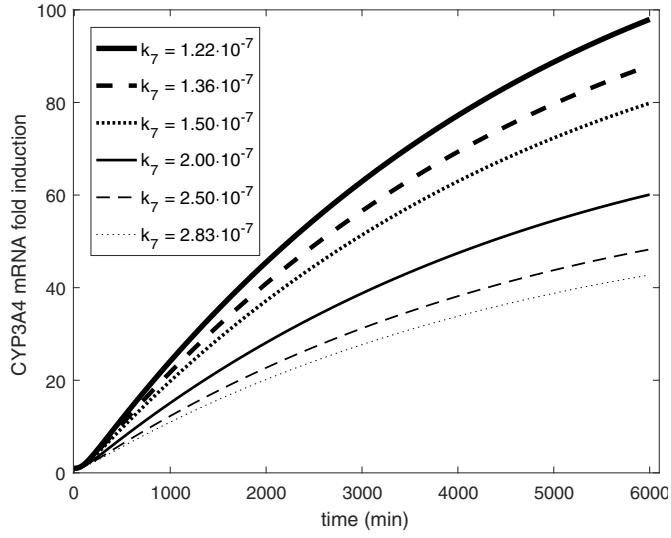

Figure 8. CYP3A4 mRNA fold induction for $s_{PXR} = 9.47{\cdot}10^{-7}$.

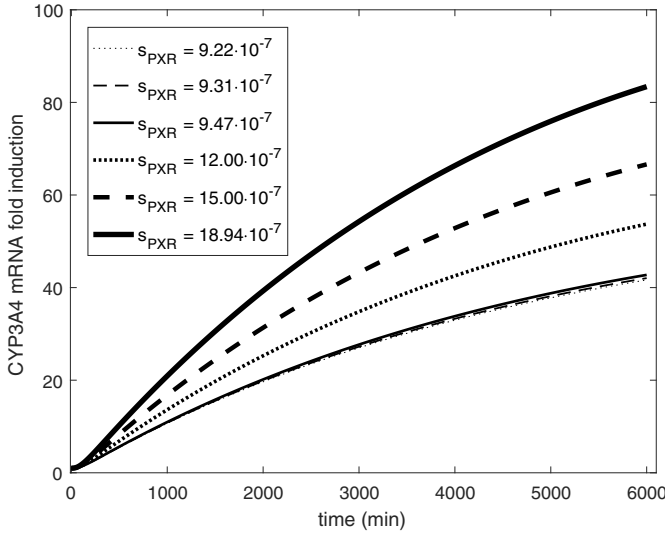

Figure 9. CYP3A4 mRNA fold induction for $k_7 = 2.83{\cdot}10^{-7}$.

Finally, we also present the parameters obtained when estimating with the standard pharmacokinetic/dynamic system software ADAPT frequently used by clinical pharmacists. In ADAPT, ODE's are solved by LSODA [16], [6], which uses variable order and variable step size formulations of Adam's method and Gear's method for respectively nonstiff and stiff equations. Function optimization for parameter estimation is done using the Nelder-Mead simplex method [14] in order to take into account the positivity constraint of the estimated pharmacokinetic/dynamic parameters. In Tables 7–8 we display some results for the simulation without or with the dosing function $d(t)$, respectively. The first two parameter columns give the estimated values obtained from minimization of the OLS function J, with the second column ignoring the apparent outlier at $t_4 = 48{\cdot}60 = 2880$ minutes. The achieved sums of squares are smaller than in the corresponding last two columns in Tables 5–6 for three out of four cases. We assume this is mainly caused by the fact that the simplex method handles better the positivity constraints for the parameters than when UFO [10] was used; in Fortran we imposed them in a naive, rather brute force way. The estimated parameter values are rather different from those generated in Tables 5–6, indicating that a different local minimum seems to have been found (though the initial guesses were always the same, except that ADAPT rounds them to the sixth digit after the comma). We remark that without the dosing function and when ignoring the outlier, ADAPT estimates the parameter s_{PXR} at roughly twice the original estimate from Luke [9] as well.

parameters	(4.1) minimal	without outlier	ML	ML without outlier
s_{PXR}	$1.26{\cdot}10^{-5}$	$1.81{\cdot}10^{-6}$	$1.61{\cdot}10^{-6}$	$1.8{\cdot}10^{-6}$
k_3	$1.95{\cdot}10^{-5}$	$1.08{\cdot}10^{-4}$	$1.82{\cdot}10^{-4}$	$1.09{\cdot}10^{-4}$
k_4	$2.51{\cdot}10^{-10}$	$7.65{\cdot}10^{-7}$	$5.34{\cdot}10^{-8}$	$7.54{\cdot}10^{-7}$
k_7	$6.69{\cdot}10^{-4}$	$5.14{\cdot}10^{-7}$	$2.9{\cdot}10^{-5}$	$1.58{\cdot}10^{-7}$
$J(k)$	116.98	9.28	216.68	9.28

Table 7. Parameter values for $d(t) = 0$ and $x_1(0) = 10\,\mu$M computed by ADAPT (ML: Maximum likelihood estimator).

parameters	(4.1) minimal	without outlier	ML	ML without outlier
s_{PXR}	$5.78{\cdot}10^{-7}$	$5.36{\cdot}10^{-7}$	$5.58{\cdot}10^{-7}$	$5.36{\cdot}10^{-7}$
k_3	$5.74{\cdot}10^{-7}$	$4.23{\cdot}10^{-7}$	$5.27{\cdot}10^{-7}$	$4.23{\cdot}10^{-7}$
k_4	$1.27{\cdot}10^{-6}$	$2.73{\cdot}10^{-6}$	$2.22{\cdot}10^{-6}$	$2.73{\cdot}10^{-6}$
k_7	$1.2{\cdot}10^{-6}$	$2.46{\cdot}10^{-6}$	$3.1{\cdot}10^{-5}$	$1.63{\cdot}10^{-6}$
$J(k)$	73.86	30.65	56.05	30.66

Table 8. Parameter values for $d(t) = 1/6$ and $x_1(0) = 0\,\mu$M computed by ADAPT (ML: Maximum likelihood estimator).

The last two columns in Tables 7–8 display the results when instead of a standard sum of squares minimization, a maximum likelihood (ML) approach is used. As expected, this does not in general give lower values for the OLS function; nevertheless, in the infusion with outlier case, ML does lead to a lower sum of squares. When ignoring the outlier, ML seems to find the same local minimum as the classical least-squares minimization. In the ML case, ADAPT estimates in addition the dependence of the measurement error on time through a linear function for this dependence. More precisely, it estimates, from user-defined initial guesses, the variances for the probability distribution of the intercept and the slope.

5. Conclusion

Most pharmacodynamic models require estimation of at least a small number of parameters. For the model of Luke et al. [9] that we investigated, we focussed on one of the four unknown parameters, the overall concentration of the nuclear receptor, because a modification of its estimate leads to a rather good fit of our own laboratory observations. While the standard sensitivity analysis presented in [9] did not predict a particularly outspoken influence of overall receptor concentration on the predicted values used for curve-fitting, we presented a different analysis showing that this influence can be approximately described by a quadratic polynomial. The proposed type of analysis might be useful for other pharmacodynamic models as well. Though

the final goal for future work, to give direct links between (1) the particular types and parameters of modelled pharmacokinetic processes and (2) properties of the Hessian matrices involved in the corresponding least-squares minimization, is out of reach for the moment, we made a first step that might contribute to its realization.

A c k n o w l e d g e m e n t. We would like to thank Prof. Petr Pávek for the laboratory experiments described in this manuscript; they were performed under his supervision and in his laboratory at the Faculty of Pharmacy of Charles university in Hradec Králové.

References

[1] *D. Z. D'Argenio, A. Schumitzky, X. Wang*: ADAPT 5 User's Guide: Pharmacokinetic/Pharmacodynamic Systems Analysis Software. Biomedical Simulations Resource, Los Angeles, 2009, Available at `https://bmsr.usc.edu/software/adapt/users-guide/`.

[2] *S. Dhillon, A. Kostrzewski,* (eds.): Clinical Pharmacokinetics. Pharmaceutical Press, London, 2006.

[3] *J. Duintjer Tebbens, M. Azar, E. Friedmann, M. Lanzendörfer, P. Pávek*: Mathematical models in the description of pregnane X receptor (PXR)-regulated cytochrome P450 enzyme induction. Int. J. Mol. Sci. *19* (2018), 1785. doi

[4] *A. Funahashi, M. Morohashi, H. Kitano, N. Tanimura*: CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. BIOSILICO *1* (2003), 159–162. doi

[5] The GNU Fortran compiler. Available at `http://gcc.gnu.org/fortran/`. sw

[6] *A. C. Hindmarsh*: Large ordinary differential equation systems and software. Control Systems Magazine *2* (1982), 24–30. doi

[7] *A. C. Hindmarsh*: ODEPACK, a systematized collection of ODE solvers. Scientific Computing 1982 (R. S. Stepleman et al., eds.). IMACS Transactions on Scientific Computation I, North-Holland Publishing, Amsterdam, 1983, pp. 55–64. MR

[8] *H. M. Jones, K. Rowland-Yeo*: Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development. CPT: Pharmacometrics & Systems Pharmacology *2* (2013), Article ID e63, 12 pages. doi

[9] *N. S. Luke, M. J. DeVito, I. Shah, H. A. El-Masri*: Development of a quantitative model of pregnane X receptor (PXR) mediated xenobiotic metabolizing enzyme induction. Bull. Math. Biol. *72* (2010), 1799–1819. zbl MR doi

[10] *L. Lukšan, M. Tůma, C. Matonoha, J. Vlček, N. Ramešová, M. Šiška, J. Hartman*: UFO 2017. Interactive System for Universal Functional Optimization. Technical Report V-1252, Institute for Computer Science CAS, Praha, 2017. Available at `http://www.cs.cas.cz/luksan/ufo.html`.

[11] *D. J. Lunn, N. Best, A. Thomas, J. Wakefield, D. J. Spiegelhalter*: Bayesian analysis of population PK/PD models: General concepts and software. J. Pharmacokinetics Pharmacodynamics *29* (2002), 271–307. doi

[12] MATLAB. Mathworks, Inc., 2018. Available at `https://www.mathworks.com/products/matlab.html`. sw

[13] *C. Moler, C. Van Loan*: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Rev. *45* (2003), 3–49. zbl MR doi

[14] *J. A. Nelder, R. Mead*: A simplex method for function minimization. Computer J. *4* (1965), 308–313. zbl MR doi

[15] NONMEM 7.3. ICON, Inc., 1990–2016. Available at `http://www.iconplc.com/innovation/nonmem/`. sw
[16] *L. Petzold*: Automatic selection of methods for solving stiff and nonstiff systems of ordinary diferential equations. SIAM J. Sci. Stat. Comput. *4* (1983), 136–148. zbl MR doi
[17] *L. Shargel, A. B. C. Yu*: Applied Biopharmaceutics & Pharmacokinetics. McGraw-Hill Education, New York, 2016.
[18] Simcyp simulator. Certara, 2012. Available at `http://www.certara.com/software/`. sw
[19] *J. W. Spruill, W. E. Wade, T. J. DiPiro, A. R. Blouin, M. J. Pruemer*: Concepts in Clinical Pharmacokinetics. American Society of Health-System Pharmacists, Bethesda, 2014.
[20] *P. Zhao, M. Rowland, S.-M. Huang*: Best practice in the use of physiologically based pharmacokinetic modeling and simulation to address clinical pharmacology regulatory questions. Clinical Pharmacology & Therapeutics *92* (2012), 17–20. doi
[21] *Z. Zheng, P. S. Stewart*: Penetration of rifampin through staphylococcus epidermidis biofilms. Antimicrob. Agents Chemoter *46* (2002), 900–903. doi

Authors' addresses: *Jurjen Duintjer Tebbens*, Faculty of Pharmacy in Hradec Králové, Charles University, Akademika Heyrovského 1203, 500 05 Hradec Králové and Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic, e-mail: `duintjertebbens@cs.cas.cz`; *Ctirad Matonoha*, Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic, e-mail: `matonoha@cs.cas.cz`; *Andreas Matthios*, Faculty of Pharmacy in Hradec Králové, Charles University, Akademika Heyrovského 1203, 500 05 Hradec Králové, Czech Republic, e-mail: `matthioa@faf.cuni.cz`; *Štěpán Papáček*, Institute of Complex Systems, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Faculty of Fisheries and Protection of Waters, University of South Bohemia in České Budějovice, Zámek 136, 373 33 Nové Hrady, Czech Republic, e-mail: `spapacek@frov.jcu.cz`.