Bayesian knowledge integration for an in vitro–in vivocorrelation model

Biometrical Journal **52** (2010) 61, zzz–zzz / DOI: 10.1002/bimj.200100000

Bayesian knowledge integration for an *in vitro–in vivo* correlation model

Elvira M. Erhardt[*,1], **Moreno Ursino** [2] , **Jeike Biewenga** [3], **Tom Jacobs** [3], and **Mauro Gasparini** [1]

[1] Department of Mathematical Sciences, Politecnico di Torino, 10129 Torino, Italy

[2] INSERM, UMRS 1138, team 22, CRC, University Paris Descartes, Sorbonne University, Paris, France

[3] Janssen Research & Development, Pharmaceutical Companies of Johnson & Johnson, 2340 Beerse, Belgium

Received zzz, revised zzz, accepted zzz

The primary goal of '*in vitro–in vivo correlation*' (IVIVC) is the reliable prediction of the *in vivo* serum concentration-time course, based on the *in vitro* drug dissolution or release profiles. IVIVC methods are particularly appropriate for formulations which are released over an extended period of time or with a lag in absorption and may support approving a change in formulation of a drug without additional bioequivalence trials in human subjects. Most of the current IVIVC models are assessed using frequentist methods, such as linear regression, based on averaged data and entail complex and potentially unstable mathematical deconvolution. The proposed IVIVC approach includes (a) a nonlinear mixed effects model for the *in vitro* release data; (b) a population pharmacokinetic (PK) compartment model for the *in vivo immediate release* (IR) data; and (c) a system of ordinal differential equations (ODEs), containing the submodels (a) and (b), which approximates and predicts the *in vivo controlled release* (CR) data. The innovation in this paper consists of splitting the parameter space between submodels (a) and (b) versus (c). Subsequently, the uncertainty on these parameters is accounted for using a Bayesian framework, i.e., estimates from the first two submodels serve as priors for the Bayesian hierarchical third submodel. As such, the Bayesian method explained ensures a natural integration and transfer of knowledge between various sources of information, balancing possible differences in sample size and parameter uncertainty of *in vitro* and *in vivo* studies. Consequently, it is a very flexible approach yielding results for a broad range of data situations. The application of the method is demonstrated for a *transdermal patch* (TD).

Key words: Bayesian; In vitro–in vivo correlation (IVIVC); Ordinary differential equation (ODE); Pharmacokinetics; Transdermal patch.

Supporting Information for this article is available from the author or on the WWW under http://dx.doi.org/10.1022/bimj.XXXXXXX

1 Introduction

Optimising the exposure of a compound becomes a key step during the drug development process. Formulation modification can occur throughout the lifetime of a compound, ranging from changes instituted during early pre-approval stages to supplementary formulation development after the initial marketing of the compound. *Controlled release* (CR) formulations, as opposed to traditional *immediate release* (IR) formulations, are a tool for such a release optimisation of the compound over time. The determination of the *in vivo* impact of a modification of the route of administration or the drug formulation is typically addressed in a clinical trial using bioequivalence testing (Chow and Liu, 2008, chap. 1.2). But studying the compound release and distribution *in vivo* through clinical trials is both time consuming and costly.
When a set of CR formulations is under development, establishing an '*in vitro–in vivo correlation*' (IVIVC)

*Corresponding author: e-mail: elvira.erhardt@polito.it

model has become an integral part of the drug development process; it links changes in the *in vitro* dissolution or release to modifications of the compound exposure. If the release remains within preexisting specifications, the exposure is assumed to remain as anticipated. For these reasons, *in vitro* drug dissolution or release can often be used as surrogate for the *in vivo* drug absorption once a, possibly nonlinear, relationship exists between the two. Under appropriate circumstances, *in vivo* bioequivalence trials may not be needed (European Agency for the Evaluation of Medicinal Products, EMEA, 1999).
The broad definition of IVIVC has been subdivided into the levels A, B, C and D. The former is considered the most informative and is consequently recommended by the American Food and Drug Administration (FDA, 1997a). It can be used to predict the entire *in vivo* time profile and, if required, the key PK parameters, e.g., *area under the curve* (AUC) or the *peak serum concentration* ($C_{\max}$).

To this end, many methods of developing IVIVC models have been proposed over the past 50 years, which generally fall into two categories: *convolution* and *deconvolution*-based methods (see e.g., the review by Yu *et al.*, 1996).
Deconvolution extracts the (unobservable) *in vivo* release based on the fact that the observable CR blood concentrations equal the convolution of IR blood concentrations and drug release *in vivo*. The latter is linked to the observed *in vitro* release results, generally via linear or nonlinear regression. Even though Dunne *et al.* (2005) and Gaynor *et al.* (2008) proved that the deconvolution method gives rise to a range of modelling and statistical problems, it is still the most frequently method used (Margolskee *et al.*, 2016). Furthermore, Gaynor *et al.* (2009) reported that averaging has a detrimental effect on the accuracy of predictions produced, still one can see quite commonly the averaging of data before its deconvolution.
The statistical limitations and presumptive mathematical instabilities associated with deconvolution-based methods explained above can be overcome using the convolution-based methods as proposed by Gillespie (1997) and O'Hara *et al.* (2001). Here, the *in vitro* release profiles and the *in vivo* blood concentration-time profiles are expressed in terms of nonlinear mixed effect models, thus, allow for the incorporation of random effects into the IVIVC estimation. Within convolution-based methods, two approaches are possible. In the *two-stage approach*, the *in vitro* release and blood concentrations from the IR formulation are modelled in a first instance. Secondly, the model predicts directly the CR blood concentration-time profiles as the convolution of the former two, avoiding explicit calculation of the *in vivo* release. In this second stage of the modelling, point estimates of the fixed and random effects from the first stage are borrowed to estimate the CR exposures. This approach splits the parameter space and model convergence, therefore, is easier obtained. However, uncertainty on the parameter estimates are not retained from one stage to another. To account for this, Jacobs *et al.* (2008) replaced the two steps of the convolution method by a *one-stage approach* through the simultaneous fitting of all three submodels. The downside of this approach is that the large parameter space in combination with the complexity of the nonlinear models might lead to slow and difficult convergence of the model fitting via maximum likelihood. Moreover, it is lacking the possibility to account for a possible imbalance of information density and study size between *in vitro* and *in vivo* study.

An intermediate solution considered in this paper is a Bayesian two-stage approach to the convolution-based method to mitigate the above mentioned pitfalls. In Bayesian analysis, prior information and observed data are combined in the context of the selected model (Gelman *et al.*, 1996). Thirty years ago, Sheiner *et al.* (1979) introduced the Bayesian pharmacokinetic approach and to our knowledge, so far merely three publications thematised Bayesian techniques in the context of level A IVIVC models. The first has been by Kortejärvi *et al.* (2006); however, their approach is restricted to very specific model and data assumptions.
Gould *et al.* (2009) developed a Bayesian IVIVC model using time-scaling to insert the release fraction into the CR model by re-scaling the time. This restricts the method to specific formulations for the IR. This method would for example fail for the commonly used intravenous formulation as IR. The objective of Qiu *et al.* (2016) was the evaluation of the correlation between *in vitro* and *in vivo* release profiles and the derivation of Bayesian tolerance intervals for the latter. A deconvolution approach is applied, ignoring the

variability among the *in vitro* repeated measurements. Further, the prediction of CR data is not addressed.

In this paper, a new Bayesian convolution-based approach to IVIVC modelling is presented that extends the current methodology. The Bayesian method was chosen since it allows for the natural integration of knowledge from one source of information into the other by combining data from several submodels in the same IVIVC model. In particular, the new aspect is the transfer of information from *in vitro* to *in vivo* setting in a flexible way, guaranteeing results even in an imbalanced data situation. The approach is illustrated on a case study, the controlled release (i.e., permeation) of a compound through a *transdermal patch* (TD), but can easily be applied to other CR routes as well. The IVIVC model was established by combining the *in vitro* release of the patch and the *in vivo* intravenous infusion blood serum concentration-time profiles of the IR formulation to predict the *in vivo* TD blood serum concentration-time profiles. The *in vitro* and IR data were first estimated using nonlinear mixed models in a frequentist approach. Subsequently, the resulting estimates served as prior information for the Bayesian hierarchical model of the TD serum concentrations. As such, the uncertainty on the fixed (population) and random (individual) effects can be integrated in the second part of the model in an intuitive manner and convergence is assured. The results are readily interpretable as probability distributions of the parameters.

This article is structured in the following way: First, the case study is introduced. In section 3, the features of the *in vitro* permeation model and the model for the *in vivo* IR formulation are explained. Thereafter, the convolution, i.e., the combination of information from the first two models in a Bayesian framework is expounded in detail. The results of applying this proposed methodology are reported in section 4. Finally, section 5 sets this new approach into perspective by offering a brief discussion and concluding remarks.

2 Case study

2.1 *In vitro* TD permeation data

The *in vitro* permeability of the compound was quantified using 50 human skin portions in order to mimic the absorption process. The skin samples were excised during surgical operations. The skin was not removed to provide samples for these *in vitro* investigations. The hospital had the prior consent of the patients that the tissue could be used for scientific research. This preclinical study has been conducted as described in Franz *et al.* (2009), which is in accordance with the guidelines of the Scientific Committee on Consumer Products (SCCP, 2006) and the Organisation for Economic Co-operation and Development (OECD, 2004a,b).
The cumulated amount (in $\mathrm{mg\,cm^{-2}}$) of drug that passed through the skin at different time-points had been recorded. Samples were taken between 0 and 72 hours after patch application. Data of two different formulations was available, i.e., a 24 h and a 72 h formulation. The $21\,\mathrm{cm^2}$ patches were loaded with $0.4\,\mathrm{mg\,cm^{-2}}$ each (72 h formulation). The 24 h formulation constituted of $0.21\,\mathrm{mg\,cm^{-2}}$ of drug in $16\,\mathrm{cm^2}$ patches. Three of the 50 skins were tested in either of the TD formulations, 47 skins only with the 72 h patch (two out of these 47 have had two patches applied). These are indicated by solid and dashed lines, respectively, in Figure 1.

The purpose of the case study in this paper is not to formally demonstrate an IVIVC relation, but rather to illustrate the proposed modelling methodology. Therefore, there was no external validation performed.

2.2 *In vivo* infusion and TD serum concentration data

Eighteen subjects had received a dose of 300 µg within 15 min via an intravenous infusion. Plasma concentration-time profiles were recorded 20 times during 36 h; they are depicted in Figure 2. Subsequently, the two transdermal patch formulations (containing 8.01 mg and 3.42 mg total dose, respectively)

were administered in eighteen different subjects in a cross-over setting, separated by a washout period. The serum concentration-time profiles are depicted in Figure 3. Further, the patches were collected after removal after 24 and 72 h, respectively, and the remaining drug substance content within the patch was assessed. The distribution of the data by patch formulation can be seen in Figure S1 in the Supporting Information on the journal's web page, which contains additional plots not presented in this manuscript.

3 Modelling

The proposed procedure consists of first fitting the *in vitro* drug permeation and the *in vivo* infusion serum concentration-time profiles. The obtained parameter estimates serve as priors for the subsequent modelling of the TD serum concentration-time profiles.

3.1 *In vitro* permeation model

The *in vitro* data collected from the skin portions served as a biomarker for the *in vivo* release of the patch. As can be seen in Figure 1, the permeation profiles described an asymmetric S-shape starting at zero. The accumulating amount of compound found in the skin portions over time can be mathematically represented by a *cumulative distribution function* (CDF). The *Weibull cumulative distribution function*,

$$\text{CDF}_W(t; h, s) = 1 - \exp\left\{-\left(\frac{t}{s}\right)^h\right\}, \qquad t \geq 0, \quad h > 0, \quad s > 0,$$

has the characteristic asymmetric S-shape and is very flexible yet not overly complex since it depends on two parameters, the shape h and the scale s. The log-normal, exponential and Gompertz CDFs were evaluated to mimic the *in vitro* data shape as well. For this purpose, the CDFs were multiplied by the total dose D_k for the two formulations $k = 1, 2$ to describe the permeation of the compound over time. However, both in the *in vitro* as well as *in vivo* setting, only a proportion of the total dose was permeated and entered the skin before patch removal, the other part remained in the patch (Figure 1 and Figure S1 in the Supporting Information), as typical for transdermal patches. Therefore, a further parameter f was introduced accounting for the *fraction* of dose delivered, i.e., the fraction of compound that was permeated from the patch.

The two formulations of the *in vitro* permeation study were fitted separately using frequentist *nonlinear mixed effects models*. This class of models is used to accommodate longitudinal data, incorporating both *fixed (population) effects* and *random (individual) effects* to account for heterogeneity between the samples. The mixed effects shape and scale had been log-transformed ($\log h = \ell h$, analogous for s) to ensure strictly positive estimates and are thus, sums of the fixed effect μ and the random effects η_i. This lead to the following model for the *in vitro* permeation measurements y_{ijk}:

$$\begin{aligned} y_{ijk} &= D_k \times f_{ik} \times \text{CDF}(t_{ijk}; h_{ik}, s_{ik}) + \epsilon_{ijk}, \quad \epsilon_{ijk} \sim \mathcal{N}(0, \sigma^2_\epsilon), \\ \ell h_{ik} &= \mu_{\ell h_k} + \eta_{\ell h,i}, \quad \ell s_{ik} = \mu_{\ell s_k} + \eta_{\ell s,i}, \qquad \eta_{\ell s,i} \sim \mathcal{N}(0, \omega^2_{\ell s_k}), \quad \eta_{\ell h,i} \sim \mathcal{N}(0, \omega^2_{\ell h_k}), \end{aligned} \tag{1}$$

where i denotes the different skin portions at $j = 1, \ldots, J_i$ measurement times t for the patch formulations $k = 1, 2$. The independent, identically, normally distributed ϵ_{ijk} refers to the measurement error, assumed to be independent of the η_i, which are normally distributed with mean 0 and variances ω^2 (Pinheiro and Bates, 2000, p. 306).

3.2 Immediate release IR serum concentrations

After the drug has been dissolved and absorbed into the bloodstream, the concentration of drug in the blood serum is determined by the bodily distribution, metabolism and elimination mechanisms (e.g., Tozer and

Rowland, 2006, chap. 2). Intravenous infusion formulations lack an absorption phase and are thus used to estimate the disposition of the compound. In this paper, the concentration of the compound in blood serum from the IR formulation was modelled by means of a compartmental model. Compartmental models represent the flow of the compound between different body tissues and are mathematically represented as a system of first-order ordinal differential equations (e.g., Rosenbaum, 2011, chap. 6.4), describing the time dependence of the drug amount in each homogeneous compartment. One- and two-compartmental models did not show a satisfactory fit to the infusion data based on a poor fit of the elimination phase as well as based on the likelihood criteria (see Figure S2 in the Supporting Information); hence, a three-compartment model was chosen for evaluation,

$$\begin{aligned}
\frac{da_1(t_{ij})}{dt} &= k_{21}a_2(t_{ij}) + k_{31}a_3(t_{ij}) - \left[k_{12}a_1(t_{ij}) + k_{13}a_1(t_{ij}) + k_e a_1(t_{ij})\right], \\
\frac{da_2(t_{ij})}{dt} &= k_{12}a_1(t_{ij}) - \left[k_{21}a_2(t_{ij})\right], \\
\frac{da_3(t_{ij})}{dt} &= k_{13}a_1(t_{ij}) - \left[k_{31}a_3(t_{ij})\right], \\
C_1(t_{ijk}) &= \frac{a_1(t_{ijk})}{V_{1,i}}.
\end{aligned} \tag{2}$$

The amount of drug in compartment $m = 1, 2, 3$ at time t is denoted by $a_m(t)$. The central compartment corresponds to the compound circulating in the blood (serum) and the compartments two and three are peripheral compartments in which the compound distributes more slowly. The parameters k_{12}, k_{21}, k_{13} and k_{31} denote the rate constants with which the drug is distributed and redistributed between central and peripheral compartments. Lastly, k_e is the elimination rate constant with which the compound is eliminated from the body. To obtain the drug concentrations C, the amounts are divided by the volume of distribution (V) of the corresponding compartments. All PK parameters have been log-transformed to avoid non-positive values and the heterogeneity between the subjects was accounted for by attributing random effects solely to $\ell V = \mu_{\ell V} + \eta_{\ell V,i}$, with $\eta_{\ell V,i} \sim \mathcal{N}(0, \omega^2_{\ell V})$, estimating all parameters with corresponding uncertainties in a nonlinear mixed effects model setting (notation as above).

3.3 Serum concentrations after transdermal patch administration

The PK of a compound is essentially describing the combined aspects of absorption and disposition. It is easy to derive that a convolution-based IVIVC modelling can be rephrased in the context of differential equations (see Buchwald, 2003). The first equation in (3) represents a complex input function (I_n), and three differential equations describe the metabolism and elimination of the compound. The *in vitro* permeation data in equation (1) can approximate the input function and the three-compartment model from the IR formulation (2) serves for describing the disposition,

$$\begin{aligned}
\mathrm{I}_n(t_{ijk}) &= D_k \times F_{ik} \times B_i \times \frac{h_{ik}}{s_{ik}} \times \left(\frac{t_{ijk}}{s_{ik}}\right)^{h_{ik}-1} \times \exp\left\{\left(-\frac{t_{ijk}}{s_{ik}}\right)^{h_{ik}}\right\}, \\
\frac{da_1(t_{ijk})}{dt} &= \mathrm{I}_n(t_{ijk}) + k_{21}a_2(t_{ijk}) + k_{31}a_3(t_{ijk}) \\
&\quad - \left[k_{12}a_1(t_{ijk}) + k_{13}a_1(t_{ijk}) + k_e a_1(t_{ijk})\right], \\
\frac{da_2(t_{ijk})}{dt} &= k_{12}a_1(t_{ijk}) - \left[k_{21}a_2(t_{ijk})\right], \\
\frac{da_3(t_{ijk})}{dt} &= k_{13}a_1(t_{ijk}) - \left[k_{31}a_3(t_{ijk})\right], \\
C_1(t_{ijk}) &= \frac{a_1(t_{ijk})}{V_{1,i}}.
\end{aligned} \tag{3}$$

The *in vitro* permeation was chosen as surrogate for the $\mathrm{I}_n(t_{ijk})$. The factor F signifies the fraction of the dose delivered (analogue to the f in equation (1)), as measured from the residual dose in patch. Part of the permeated amount of the dose potentially did not enter the bloodstream; hence, it did not contribute to the serum concentration-time profile, implying the need for a biovailability factor B, the fraction of dose delivered which is actually absorbed into the systemic circulation. To ensure it is laying within the unit interval after inclusion of the random effects, a logit-transform was applied to B, so that the transformed $\tilde{B}$ can range between $\pm\infty$. The bioavailability had assigned a random effect per subject i; the same held for the PK parameter $V_{1,i}$, in line with the infusion model (2).
The vector θ unifies all twelve population parameters estimated by the model,

$$\theta = (\mu_{\tilde{B}}, \mu_{\ell h_k}, \mu_{\ell s_k}, \mu_{\ell V}, \mu_{\ell k_{12}}, \mu_{\ell k_{21}}, \mu_{\ell k_{13}}, \mu_{\ell k_{31}}, \sigma^2), \qquad k = 1, 2,$$

and the standard deviations (SDs) of the four random effects are

$$\Omega = (\omega_{\tilde{B}}, \omega_{\ell h}, \omega_{\ell s}, \omega_{\ell V}).$$

To restrict the model's complexity to a (computationally) reasonable extent, the mixed effects ℓh_{ik} and ℓs_{ik} for the dose-dependent shape and scale, respectively, of the *in vitro* permeation model's input function (1) had their specific fixed μ_k but shared the random η_i,

$$\theta_{n,ik} = \mu_{n,k} + \eta_{n,i} \qquad \text{with} \qquad \eta_{n,i} \sim \mathcal{N}(0, \Omega_n^2), \qquad n = 1, 2, 3.$$

The index n is denoting the nth element of the vector θ and Ω, respectively. The small sample size of the 24 h formulation in the permeation data (see the three dashed lines in Figure 1) yielded an artificially low variability of its parameter estimates. To be conservative, the higher estimates of the 72 h formulation have been used for the shared random effects of ℓh and ℓs.

An additive residual error (i.e., assuming normal distribution) or a proportional residual error (i.e., assuming log-normal distribution) model is common practice when modelling serum concentration-time profiles. Instead of using transformations to obtain normally distributed residuals, one can change the distributional assumption itself. Lindsey *et al.* (2001) showed that the gamma distribution is often more suitable when fitting PK concentration data. As this held true for the TD data in the present case, this recommendation has been followed. Accordingly, the TD observations $y(t_{ijk})$, i.e., the measured drug concentrations in blood serum of the TD formulation in subject i at time t_j wearing patch formulation k were modelled as:

$$\begin{aligned} y(t_{ijk}) &\sim \mathcal{G}(\alpha(t_{ijk}), \beta(t_{ijk})), \\ \alpha(t_{ijk}) &= (C_1(t_{ijk})/\sigma)^2, \\ \beta(t_{ijk}) &= C_1(t_{ijk})/\sigma^2. \end{aligned}$$

The shape α and the scale β parameter of the gamma distribution have been defined in order to maintain the mean value, $\mathrm{E}(y(t_{ijk})) = C_1(t_{ijk})$, and the variance, $\mathrm{Var}(y(t_{ijk})) = \sigma^2$, of the observed TD data under the assumed model.

3.4 Bayesian knowledge integration

The integration of component knowledge to a whole system requires a systematic and yet flexible approach. The Bayesian algorithm enables a natural flow of parameter uncertainty between the different models where parameter estimates of the first stage serve as priors for the second stage of the modelling.
A characteristic difficulty of PK model estimation is that they predict a pattern of concentration over time that is close to a mixture of declining exponential functions, with the amplitudes and decay times of the different components corresponding to functions of the model parameters. The estimation of the

decay times of a mixture of exponentials is known to be an ill-conditioned problem (see e.g., Acton, 1970, p. 253), which results in difficulties to simultaneously estimate the parameters in such a model (Gelman *et al.*, 1996). Furthermore, the simultaneous estimation of twelve fixed effects plus four random effects is of great complexity, above all from observations of only 36 patch applications. These circumstances, as well as the availability of relevant information on many parameters, justify the choice of informative prior distributions with, in some cases, low SDs (see Table 1). Further, for the sake of simplicity, all parameters were assumed to be independent and thus, only univariate priors had been used.
The log-transformation of PK and Weibull parameters yielded normal distributed priors; to be coherent, a weakly informative half-normal distribution was assigned to the residual variability σ^2 of the TD data. The following joint prior resulted,

$$p(\theta,\Omega) = p(\theta) \times p(\Omega) = \left[\left\{\prod_{n=1}^{11} \mathcal{N}(\hat{\mu}_n, \hat{\sigma}_n^2)\right\} \times \text{half-}\mathcal{N}(0,\delta)\right] \times \left[\prod_{n=1}^{4} \text{half-}\mathcal{N}(\hat{\omega}_n, \xi_n^2)\right],$$

with specific means $\hat{\mu}_n, \hat{\omega}_n$ and variances $\hat{\sigma}_n^2$ equal to the estimates from the IR and permeation datasets and corresponding standard effects, respectively (cf. Table 1). The half-normal distribution scale parameter $\delta = 10$ and the variances $\xi_n^2 = 2$ of the elements in Ω were chosen high to remain conservative. For the same reason, the prior SD of the parameters $\mu_{\ell s_k}$ was increased to 1. Further, independence was assumed for model parameters and variance parameters, although a multivariate approach would have been more suited. It was abstained from this to facilitate the interpretability of the model parameters and to evade highly intricate structures.

To evaluate the model ability of predicting the correct serum concentration data, *population predictions* were simulated. That is, the model was used to repeatedly simulate observations according to the original design of the study, resulting in predictions of future observations $\tilde{y}$ for additional subjects (exchangeable with the subjects in the study). This was achieved by simulating random vectors of the physiological parameters from their simulated population distribution. The posterior predictive distribution was calculated as

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta,\Omega) p(\theta,\Omega|y) d(\theta,\Omega).$$

The variance in the population distribution of the parameters includes posterior uncertainty in the parameter estimates and also real variation in the population.

To compare directly the impact of the integration of information via prior uncertainty, a version of the proposed CR model *without uncertainty propagation* has been applied as well. That is, the point estimates from the first stage (*in vitro* and IV model) were plugged in as fixed data and only the bioavailability B and the error term σ had priors attached and thus, were estimated. Then, the predictions were conducted analogously to the way described above.

Key quantities of interest in concentration prediction, the AUC, the $C_{\max}$ and the *time of peak concentration*, $t_{\max}$, were calculated for model evaluation. Those were compared between observations and predictions and between the proposed model with uncertainty and the and version without uncertainty propagation. At last, the *percent prediction error*

$$\%PE = 100 * \frac{|AUC_{\text{observations}} - AUC_{\text{predictions}}|}{AUC_{\text{observations}}}$$

was calculated as well as a measure of prediction accuracy.

4 Results

The two frequentist nonlinear mixed effects models for the *in vitro* release formulations were fitted using the R package `nlme` by Pinheiro and Bates (2000, chap. 8.2). The function approximated the *in vitro* data well. It has been formally compared with other distributions based on the Akaike Information Criterion (AIC). For both formulations, the log-normal distribution yielded the lowest AIC. Looking at the 72 h formulation, the result was 8664.08 for the log-normal, followed by the Weibull (8672.22), the Gompertz (8821.10) and lastly, the exponential distribution (9106.94). The decision has been made for the Weibull CDF, considering the small difference in AIC between its fit and the log-normal fit, the visual difficult distinguishability of their curves (Supporting Information, Figure S3), the common usage of the Weibull distribution in drug release curve fitting (Goldsmith *et al.*, 1978; Costa and Lobo, 2001; Zhou, 2003) and the interpretability of its shape parameter as permeation rate in that context.
The graphical fit of the permeation profiles using only the population effects for shape (h) and scale (s) can be found in Figure 1. Figure S4 in the Supporting Information depicts the individual fits to skins using the random effects η for h and s; the parameter estimates are in Table 1 (first column, top).
For the nonlinear mixed effects model of the IR concentration-time profiles, the `nlmixr` package in R was utilised. The adequate model fit of the three-compartment model for the IR data is shown graphically in Figure 2 (population fit) and in the first column (bottom) of Table 1. A plot of the individual fit to the infusion data is given in the Supporting Information Figure S5. The figures show a very good approximation of the data, except for one observation of subject number 111.
The TD data was fitted using the combined ODE system (3) of the drug permeation model (1) and the infusion model (2) above. The frequentist parameter estimates (with uncertainties) of these two models were utilised as initial values and prior information for the Bayesian fit of the IVIVC model. The resulting posterior distributions involve high-dimensional integration and are analytically intractable. This hierarchical Bayesian differential equation modelling was implemented via *Hamiltonian Monte Carlo* (HMC) (e.g., Neal , 2011), a type of *Markov Chain Monte Carlo* (MCMC) algorithm incorporated in the software package `Stan`, via `R` (Stan Development Team , 2018). To make the MCMC method arithmetically more stable and to improve the sampling efficiency of the HMC algorithm, all mixed effects were reparametrised non-centrally (Betancourt and Girolami, 2015; Papaspiliopoulos *et al.*, 2007). `Stan`'s HMC algorithm was used to draw samples of the parameters from the posterior distribution via the R package `rstan`. From those, features of the marginal distributions of interest could be inferred as well as model diagnostics such as visual predictive checks (VPCs).
To allow faster convergence, only the shapes h and scales s of the Weibull distributions, the volume V from the PK parameters and the bioavailability B had random effects added. The volume was estimated for each of the 18 subjects, the other random effects for each of the 36 patch applications. Both patch formulations were fitted simultaneously to allow an exchange of information between formulations.
The individual predictions (per subject) of the TD data can be seen in Figure 4. More precisely, the real observations with the posterior median and the 90% credibility interval (CI) of the simulated observations are depicted. These are predictions of newly observed concentrations in the same subjects participating in the actual study. Figure 5 contains the 90% CI of the population predictions, the predicted new observations in new subjects. For an improved illustration of the model's predictive ability and appropriateness, quantiles of the entirety of population predictions have been plotted versus time on top of the corresponding quantiles of the real observations, to aid comparison of predictions with observations (VPC). Ideally, the solid curves are congruent with the dashed lines or at least, lay within the 50% confidence band around each of the prediction intervals obtained by simulation (grey areas in the plot). The percentiles chosen for plotting were the 5^{th}, 50^{th} and 95^{th} percentile (Figure S6 in the Supporting Information) as well as the 25^{th}, 50^{th} and 75^{th} (Figure S7 in the Supporting Information) as for small data sets, less extreme percentiles may be more appropriate (Karlsson and Holford, 2008).
The comparison of the AUCs between observed and individual predicted concentrations for the models with and without uncertainty propagation can be seen in Figure 6. The densities of the corresponding

$\%PE$s with the quantiles are overlaid as vertical red bars are plotted as well (see Figure 7). Further key quantities and comparison also with the population predictions are contained in the Supporting Information (Figures S8-S15).

For each parameter, 4 independent Markov chains were started at values obtained by sampling the corresponding parameter at random from its prior distribution. After careful inspection of former programme runs, the number of iterations per chain had been fixed to 750, of which the first 150 were discarded as "warm-up", adding up to 2400 iterations in total. Compared to iteration numbers of MCMC algorithms this seems few but it is important to note that the HMC algorithm explores the parameter space much more efficiently. For the population predictions, 2400 sets of 18 new subjects receiving both formulations were simulated, yielding $540 \times 2400 = 1,296,000$ predicted observations. The computations required ca. 30 hours on a Linux Infiniband-QDR MIMD Distributed Shared-Memory Cluster with CPU Model 2x Xeon E5-2680 v3 2.50 GHz (turbo 3.3 GHz) 12 cores.
The convergence of the simulations was monitored by comparison of the variance between and within sequences for all parameters of interest. Specifically, the $\hat{R}$ statistic by Gelman and Rubin (1992) was calculated to assess convergence comparing within- and between-chain variability over the second half of each chain (see Table 1 and Figure S16 in the Supporting Information). Additionally, the chain's traces (the time series of the posterior draws), the pairwise correlations and HMC diagnostics and each parameter's Kernel density of posterior draws were plotted to assess convergence visually (see the Supporting Information Figures S17-S18). The chain's plotted traces resembled "fuzzy caterpillars", a sign for convergence, and there were 0 divergences after warm-up reported by the Stan algorithm. The plot of the parameter's pairwise correlations was free of abnormalities or strong dependencies.
Source code and simulated datasets to reproduce the results are available among the Supporting Information on the journal's web page.

5 Discussion

An *in vitro–in vivo* correlation is a mathematical description of the, possibly nonlinear, relationship between *in vitro* drug release and either *in vivo* drug release or absorption. The IVIVC models are commonly used in a wide range of applications. During clinical phases I and II an IVIVC model can be used for modification or optimisation of the drug formulation in order to achieve target concentrations of drug in serum which remain within the therapeutic window over a sufficient period of time. In clinical phase III and in the post approval period, an IVIVC model can support changes made to production procedures, known as variations or scale-up and post approval changes (SUPAC) (FDA, 1997b; Leeson, 1995) and cost-intensive *in vivo* bioequivalence trials might be evitable.
IVIVC modelling consists of combining of data from different resources in a physiological meaningful manner. Several methods have been proposed. The most widely applied among these are deconvolution-based. This may be due to their relatively easy implementation. Gaynor *et al.* (2008) demonstrated that bias might be introduced in this method. Therefore, convolution-based methods seem more appropriate for the development of IVIVC modelling, despite the more challenging implementation of such models.
The goal of this project was to investigate a fundamental way of integrating prior knowledge and TD data to obtain a full picture of the IVIVC. A Bayesian nonlinear mixed-effects IVIVC model is introduced and the method is applied to the case study of a transdermal patch. The hierarchical aspect is deemed required to correctly account for the sources of variation, i.e., inter- versus intra-subject variability. This aspect is often ignored in deconvolution–based methods, where the average *in vitro* release profile is used. In the two-stage convolution-based methods, the parameter uncertainty is also not accounted for when parameter estimates are transferred to the second stage, instead it is claimed that the first stage estimates are estimated exactly and all uncertainty is therefore pushed forward to the second stage in the estimation process, which potentially leads to inflated variances. On the other hand, the one-stage modelling might lead to

a compensation-type of effect from one submodel towards the other submodels. More detailed, in practice there is often an imbalance of sample size between *in vitro* and *in vivo* studies, typically the latter is sparser due to its higher costs, but also the contrary, very rich clinical trials, can occur. In this situation, the one-stage approach would tend to discard the information of the smaller trial, as the larger trial might dominate the estimation process. In the pharmacokinetic/pharmacodynamic (PKPD) modelling context, simultaneous modelling of different sources of data is statistically preferable compared to first estimating parameters, which are then fixed and used in the second stage of the analysis (Zhang *et al.*, 2003). Jacob *et al.* (2017) investigate in their paper why and when modular approaches might be preferable to the full model and propose a method for choosing between these two types of inference.
The applicability of Bayesian statistics in IVIVC models has been recognised about 10 years ago by Kortejärvi *et al.* (2006). The limiting assumptions of their approach, however, are that oral dosage form absorption is dissolution rate controlled, pharmacokinetics follows a one-compartment model with first order absorption and the AUC is independent of the formulations. Moreover, the dissolution function was fit to averaged dissolution data. The posterior parameter estimates were generated in absence of an IR unit, thus, the method works only with very informative CR data and a simple compartment model to reduce the number of parameters. Gould *et al.* (2009) developed a Bayesian predictive approach for IVIVC which rescales the time based on the *in vitro* release so that the *in vivo* release of the IR corresponds to it. Therefore, the method can not be used with an intravenous infusion as IR. Qiu *et al.* (2016) proposed a method to derive two-sided Bayesian tolerance intervals depending upon averaged observed *in vitro* release and deconvolved *in vivo* release profiles.

A Bayesian hierarchical model is a powerful and natural tool for incorporating such parameter uncertainty between different stages of the modelling. The approach explained in this paper represents an intermediate solution to the possible criticism on both the one- and the two-stage convolution-based methods. It uses the estimated means and corresponding SDs from the first stage – the fit of the *in vitro* permeation and the *in vivo* IR data – merely as prior information for the second stage. Thus, the (lack of) information on parameters is shared towards modelling of the CR formulation rather than using plug-in estimates from the first stage of PK modelling which may introduces an undue over-confidence into the model.
By following this approach, preclinical *in vitro* and *in vivo* IR data is used not only to extract the parameter estimates, but also to understand the CR model structure. Once the CR model is built in this way (including some transformation, if necessary) the before acquired knowledge about the parameters of two separate sources enters the model together as prior information to predict the exposure of the new formulation in human blood serum. This step-by-step procedure is a general approach that can come flexibly into operation for many different kinds of dissolution or release, IR and CR studies. The method is also feasible in a very imbalanced data situation with high uncertainty around the parameters, since it uses the first stage data to gather knowledge but stays flexible by usage of informative priors. Consequently, convergence is assured and uncertainty is accounted for in a more intuitive manner, i.e., both in the *in vitro* and the *in vivo* part of the IVIVC model.

When applying this approach to the case study at hand, the resulting Figures 1 and 2 show that the population effects of the *in vitro* permeation and the *in vivo* IR serum concentration-time profiles, respectively, are laying well within the individual curves. The model fit was evaluated based on likelihood criteria as well as visual inspection. The appropriate fit to the *in vitro* permeation data could be expected as the Weibull CDF has a long tradition in the modelling of release profiles (Christensen *et al.*, 1980; Berry and Likar, 2007), with Piotrovskii (1987) being the first one applying it, to our knowledge.
The individual fits are tracing the observations of single skin and subject, respectively (see the Supporting Information Figures S4 and S5, respectively, for the graphical representation). The only remarkable deviation from the measurements is the fifth sample of subject number 111 in Figure S5, which might be considered as a potential outlier in Figure 2.

The individual predictions for the TD serum concentration-time profiles, shown in Figure 4, fit the observations well. It is important to note that in this study, the *in vitro* permeation (in skin portions) has been set equal to the *in vivo* permeation in subjects ($\mathrm{I}_n(t_{ijk})$). This is generally not the case and we acknowledge that this is an unrealistic assumption. However, as the case study serves for illustrative purposes only, it was analysed in a basic manner. Generalisation can be done easily as proposed by Dunne *et al.* (1999), i.e., the *in vivo* release would be replaced by the *in vitro* release upon generalisation with a link function.
Solely the absorption phase of the 24 h patch worn by subjects number 1017 and 1018 in Figure 4 were not fitted optimally by the posterior median. The late onset of the absorption phase of subject 1017 is difficult to fit as well as the initial burst of the latter. Figure 5 suggests that the predictions made by the model are in agreement with the data. The sharp discontinuity of the posterior median and credibility bands at 72 and 24 h, respectively, is due to the removal of the patch from the skin, after which no compound was absorbed any more.
The VPC reveals that the observation median lies for the most part within the 50% confidence bands of the mean posterior prediction, the lower and upper percentiles instead, are estimated too low and too high, respectively. This discrepancy is apparent for both patch formulations but more distinct for the second (lower dose).
From the comparison of the AUCs (Figures 6 and 7) and other key quantities (Figures S8-S15 in the Supporting Information) for the different scenarios can the conclusion be drawn, that the propagation of uncertainty results in an advantageous prediction, thus, the estimation in the Bayesian framework brings a clear improvement of the IVIVC model. Only for the AUC of the 72 h formulation, the optimal result of a straight line in the comparison plots could be achieved and the small sample size of the 24 h *in vitro* formulation gives in general less accurate predictions than the 72 h formulation. Accordingly, the $\%PE$s lie above the FDA criteria of 10% or less for AUC and $C_{\max}$ to establish the predictability of the IVIVC (FDA, 1997a). However, the aim of this work is not the claim of establishment of an IVIVC for the compound according regulatory guidelines but rather, to outline a general framework of a new IVIVC development methodology.
According to (Källén, 2007, p. 23), the bioavailability B is not only subject to between-subject variability, but it can also vary between different doses k. However, a model with random effect assigned per subject i and for each patch formulation k lead to a strong positive correlation of the two biovailability factors. Reparametrisation removed the correlation and revealed a 98% congruousness of B_1 and B_2, wherefore the model with a single dose-independent B was chosen.
In general, when first stage results are used for building priors of a second stage, this can possibly result in inconsistent estimates as prior-data conflicts can arise. The IVIVC theory however, is based on the concept that the *in vitro* information surrogates the *in vivo* environment. Hence, a conflicting prior would indicate that one cannot make any conclusions regarding the *in vivo* data based on the *in vitro* release profiles. Therefore, such data would support absence of an IVIVC model.
Still, to avoid overconfidence into the frequentist estimates, a sensitivity analysis with priors having SDs twice as high (with the exception of the already high Ω SDs) was run. As can be seen in Table S1 in the Supporting Information, the posterior means are comparable, while some of the posterior SDs are increased.
Given the vastly increased run times (by a factor of 3.5 from ca. 30 to ca. 104 hours) caused by the broader parameter space, the good practice was followed and not deviated from the frequentist variance estimates of the first stage.
Lastly, a gamma distribution has been proposed for the serum plasma concentrations and was straightforwardly implemented in `Stan`, while commonly distributions other than the (log-)normal distribution so far remain underused, particularly due to their unavailability in standard PK software (Dosne *et al.*, 2016).

In conclusion, the presented combined *in vitro–in vivo* model provides a satisfactory estimation of the transdermal patch serum concentration-time profile. The Bayesian framework allows a natural integration of knowledge from one source of information into another, while accounting for the uncertainty of the

parameters. The possibility to utilise all observed data when constructing an IVIVC model, can be considered as a major strength of the Bayesian approach. This is a clear improvement of the current IVIVC methodology where a frequentist one-stage or two-stage approach is the standard. The proposed IVIVC model enables the use of *in vitro* release for formulation optimisation and can function as a surrogate for *in vivo* bioequivalency studies of the transdermal patch at hand.
Further, the proposed approach is very flexible: it can be easily modified to combine multiple sets of IVIVC data with relevant sources of variation incorporated into the analysis. Also, the ODE system makes a switch to formulations following Michaelis-Menten kinetics (as opposed to linear elimination) straightforward.

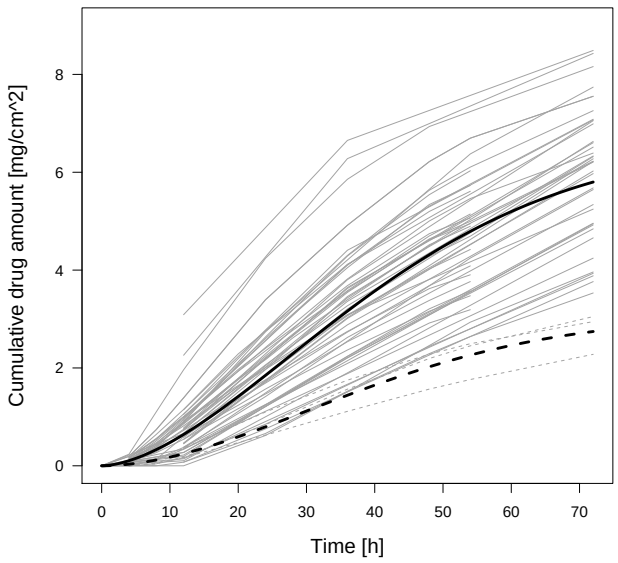

Figure 1 Population fit of the Weibull input functions (black) to the 52 permeation curves of the 72 h formulation (grey, solid) and the 3 permeation curves of the 24 h formulation (grey, dashed).

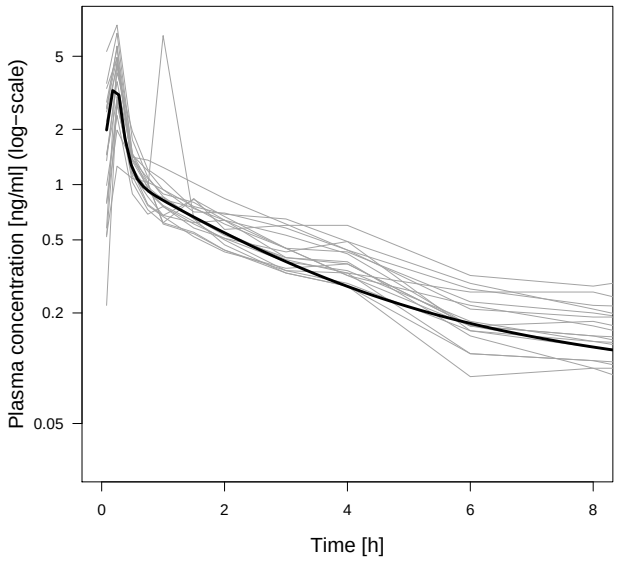

Figure 2 Population fit (black) of the compartment model to the 18 IR profiles (grey).

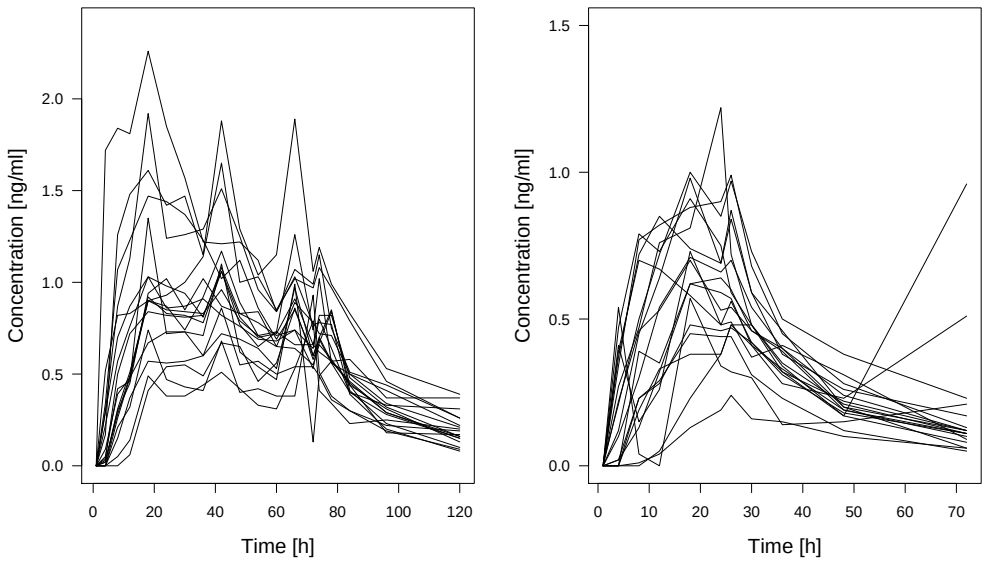

Figure 3 Serum concentrations of the TD's 72 h formulation (left) and 24 h formulation (right), 18 concentration-time profiles each.

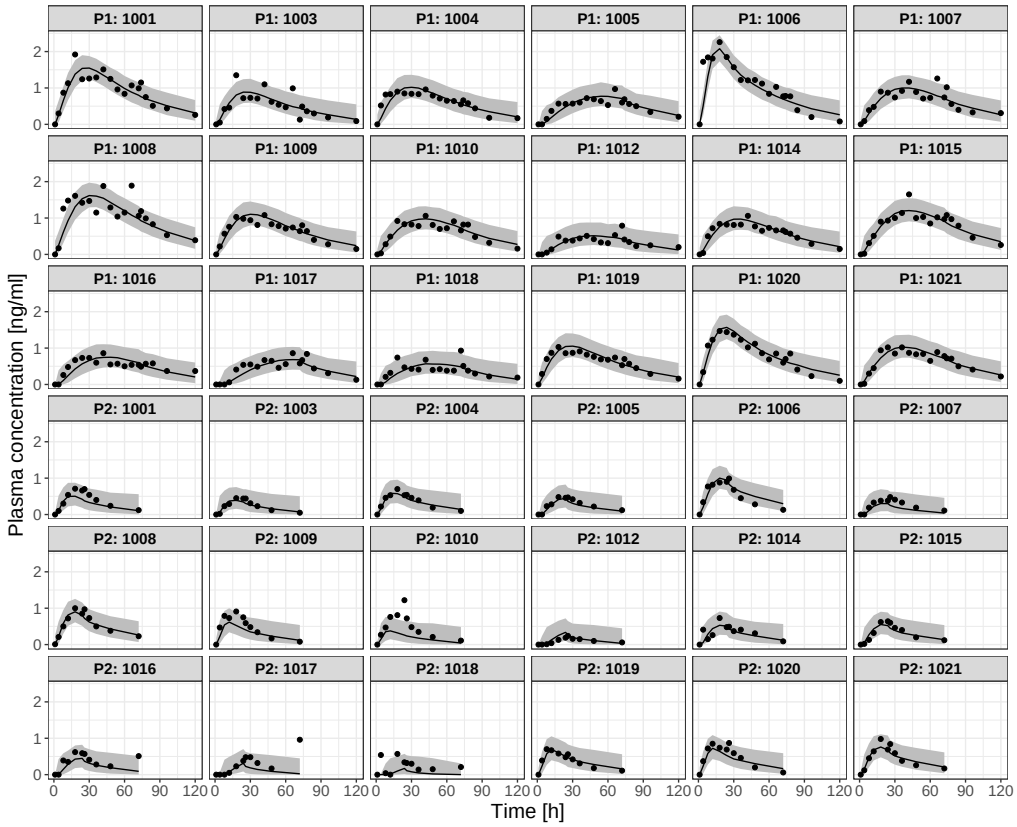

Figure 4 Serum concentrations (black dots) with posterior median (black lines) and 90% credible bands of predicted individual observations in study subjects.

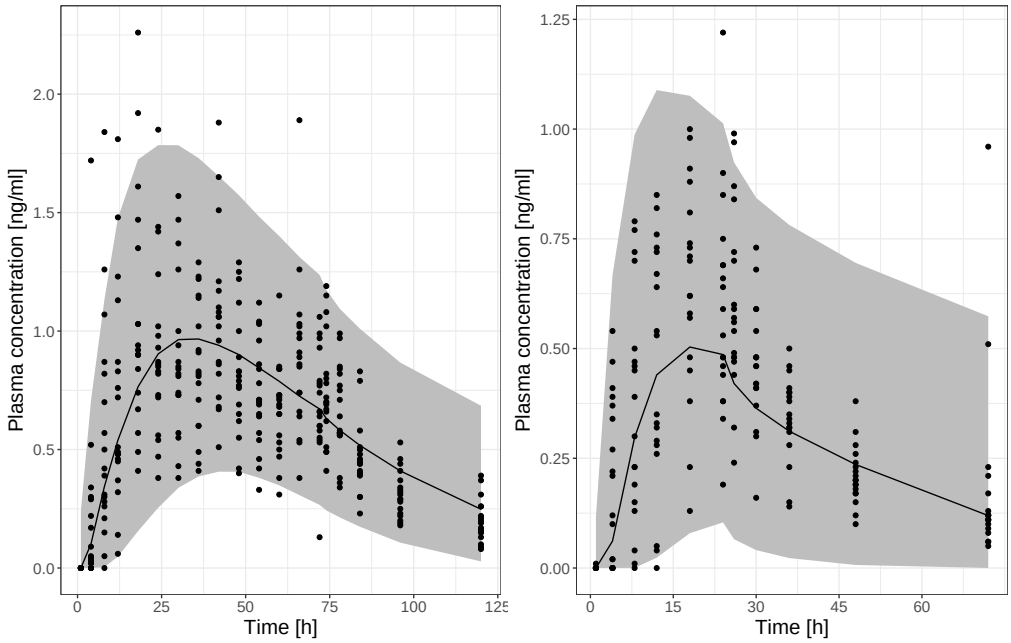

Figure 5 Serum concentrations (black dots) with posterior median (black lines) and 90% credible bands of predicted population observations in new subjects for the 72 h formulation (left) and the 24 h formulation (right) of the TD.

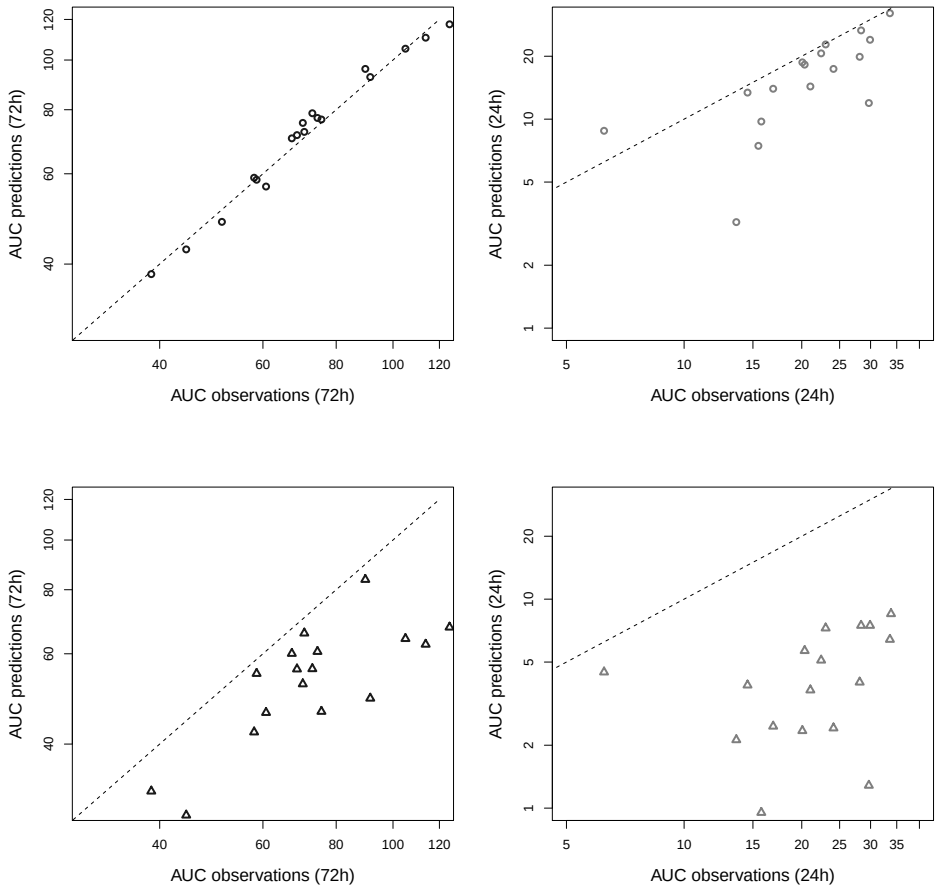

Figure 6 Individual predictions: observed vs. predicted AUC for the 72 h formulation (left column, dark grey) and the 24 h formulation (right column, light grey) comparing the proposed model (top row, circles) with the model without uncertainty propagation (bottom row, triangles). The dashed line indicates the diagonal.

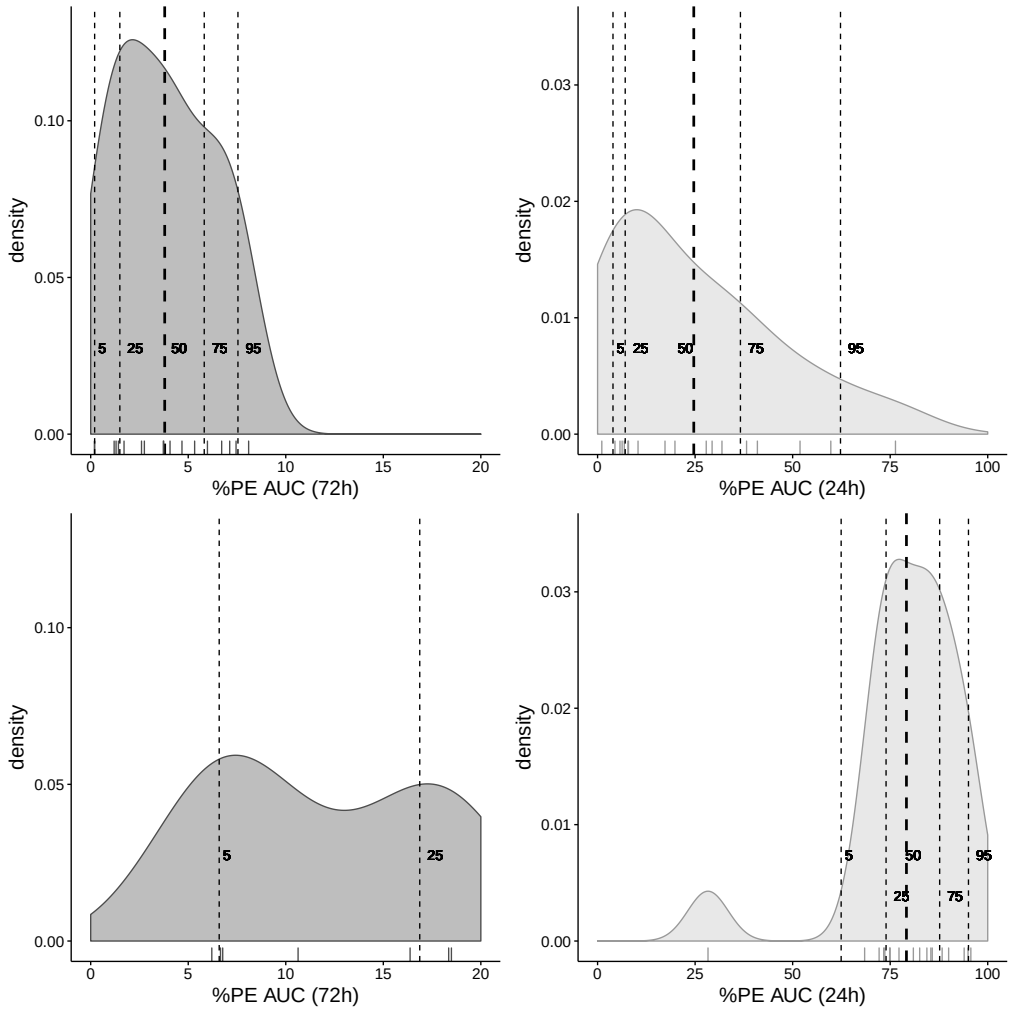

Figure 7 Individual predictions: density of the AUC percent prediction errors ($\%PE$) for the 72 h formulation (left column, dark grey) and the 24 h formulation (right column, light grey) comparing the proposed model (top row) with the model without uncertainty propagation (bottom row). Quantiles are overlaid as vertical dashed lines.

Table 1 Estimates of all model parameters. Frequentist models for drug permeation and IR with fixed effects, corresponding SDs and random effect SDs (both in parenthesis). Analogously for the Bayesian model of the TD data, with posterior SD (in parenthesis) and Gelman-Rubin statistic $\hat{R}$.

Parameter	mean (SD)					
	Frequentist		**Bayesian**			
	fixed	random	fixed	$\hat{R}$	random	$\hat{R}$
	permeation		*controlled release*			
B	-	-	0.34 (0.09)	1.01	0.26 (0.15)	1.03
f_1	0.83 (0.02)	(0.16)	-	-	-	-
f_2	0.91 (0.05)	(0.07)	-	-	-	-
$\log(h_1)$	0.52 (0.02)	(0.15)	0.52 (0.02)	1.00	0.08 (0.02)	1.00
$\log(h_2)$	0.61 (0.05)	(0.05)	0.66 (0.03)	1.00		
$\log(s_1)$	3.84 (0.03)	(0.16)	3.34 (0.13)	1.00	0.42 (0.06)	1.00
$\log(s_2)$	3.82 (0.07)	(0.11)	2.72 (0.12)	1.01		
	immediate release					
$\log(V)$	10.47 (0.11)	(0.38)	10.50 (0.11)	1.00	0.17 (0.09)	1.00
$\log(k_e)$	0.39 (0.28)	-	-1.00 (0.19)	1.01	-	-
$\log(k_{12})$	0.15 (0.45)	-	0.32 (0.42)	1.00	-	-
$\log(k_{21})$	-2.16 (1.05)	-	-2.12 (0.28)	1.00	-	-
$\log(k_{13})$	1.82 (0.15)	-	2.13 (0.15)	1.00	-	-
$\log(k_{31})$	0.61 (0.20)	-	0.35 (0.17)	1.00	-	-
σ	-	-	0.18 (0.01)	1.00	-	-

Acknowledgements Elvira Erhardt, Tom Jacobs and Mauro Gasparini have received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567. Moreno Ursino was funded by the French National Cancer Institute (INCa) grant INCA_9539. Further, Janssen Pharmaceutica is acknowledged for providing the data. Computational resources were provided by HPC@POLITO (`http://hpc.polito.it`).

Authors' contributions

Tom Jacobs and Mauro Gasparini made equal contributions.

Conflict of Interest

The authors have declared no conflict of interest.

References

Acton, F. S. (1970). Numerical methods that work. *Harper and Row, New York* .

Berry, M. R. and Likar, M. D. (2007). Statistical assessment of dissolution and drug release profile similarity using a model-dependent approach. *Journal of Pharmaceutical and Biomedical Analysis* **45**, 194 – 200.

Betancourt, M. and Girolami, M. (2015). Hamiltonian Monte Carlo for hierarchical models. *Current Trends in Bayesian Methodology with Applications* **79**, 30.

Buchwald, P. (2003). Direct, differential-equation-based in-vitro-in-vivo correlation (IVIVC) method. *Journal of Pharmacy and Pharmacology* **55**, 495—504.

Chow, S.-C. and Liu, J.-P. (2008). *Design and Analysis of Bioavailability and Bioequivalence Studies*. CRC press.

Christensen, F. N., Hansen, F. Y., and Bechgaard, H. (1980). Physical interpretation of parameters in the Rosin-Rammler-Sperling-Weibull distribution for drug release from controlled release dosage forms. *Journal of Pharmacy and Pharmacology* **32**, 580–582.

Costa, P. and Lobo, J. M. S. (2001). Modeling and comparison of dissolution profiles. *European Journal of Pharmaceutical Sciences* **13**, 123 – 133.

Dosne, A., Bergstrand, M., and Karlsson, M. (2005). A strategy for residual error modeling incorporating scedasticity of variance and distribution shape. *Journal of Pharmacokinetics and Pharmacodynamics* **43**, 137–151.

Dunne, A., Gaynor, C., and Davis, J. (2005). Deconvolution Based Approach for Level A In Vivo-In Vitro Correlation Modelling: Statistical Considerations. *Clinical Research and Regulatory Affairs* **22**, 1–14.

Dunne, A., O'Hara, T., and Devane, J. (1999). A new approach to modelling the relationship between in vitro and in vivo drug dissolution/absorption. *Statistics in Medicine* **18**, 1865–1876.

EMEA (1999). Note for Guidance on Quality of Modified Release Products: A: Oral Dosage Forms B: Transdermal Dosage Forms Section I (Quality). https://goo.gl/3EXvGW.

FDA (1997a). Guidance for Industry: Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations. http://www.gmp-compliance.org/guidemgr/files/1306FNL.pdf.

FDA (1997b). Guidance for Industry, SUPAC-MR. 1997. Modified-release solid oral dosage forms. https://www.fda.gov/downloads/Drugs/Guidances/ucm070640.pdf.

Franz, T. J., Lehman, P. A. and Raney, S. G. (2009). Use of excised human skin to assess the bioequivalence of topical products. *Skin Pharmacology and Physiology* **22**, 276 – 286.

Gaynor, C., Dunne, A., and Davis, J. (2008). A Comparison of the Prediction Accuracy of Two IVIVC Modelling Techniques. *Journal of Pharmaceutical Sciences* **97**, 3422 – 3432.

Gaynor, C., Dunne, A., and Davis, J. (2009). The effects of averaging on accuracy of IVIVC model predictions. *Journal of Pharmaceutical Sciences* **98**, 3829–3838.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Journal of the American Statistical Association* **1**, 515–534.

Gelman, A., Bois, F., and Jiang, J. (1996). Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions. *Journal of the American Statistical Association* **91**, 1400–1412.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7**, 457–472.

Gillespie, W. (1997). Convolution-based approaches for in vivo-in vitro correlation modeling. *Advances in Experimental Medicine and Biology* **423**, 53–65.

Goldsmith, J. A., Randall, N., and Ross, S. D. (1978). On methods of expressing dissolution rate data. *Journal of Pharmacy and Pharmacology* **30**, 347–349.

Gould, A. L., Agrawal, N. G. B., Goel, T. V., and Fitzpatrick, S. (2009). A 1-step Bayesian predictive approach for evaluating in vitro in vivo correlation (IVIVC). *Biopharmaceutics & Drug Disposition* **30**, 366–388.

Jacob, P. E., Murray, L. M., Holmes, C. C. and Robert, C. P. (2017). Better together? Statistical learning in models made of modules. *arXiv preprint arXiv:1708.08719.* https://arxiv.org/abs/1708.08719.

Jacobs, T., Rossenu, S., Dunne, A., Molenberghs, G., Straetemans, R., *et al.* (2008). Combined Models for Data from In Vitro–In Vivo Correlation Experiments. *Journal of Biopharmaceutical Statistics* **18**, 1197–1211.

Källén, A. (2007). *Computational Pharmacokinetics*. Chapman & Hall/CRC Biostatistics Series. Taylor & Francis.

Karlsson, M. O. and Holford, N. (2008). A Tutorial on Visual Predictive Checks. *Talk presented at the Annual Meeting of the Population Approach Group in Europe (PAGE).* https://www.page-meeting.org/?abstract=1434.

Kortejärvi, H., Malkki, J., Marvola, M., Urtti, A., Yliperttula, M., *et al.* (2006). Level A in vitro-in vivo Correlation (IVIVC) Model with Bayesian Approach to Formulation Series. *Journal of Pharmaceutical Sciences* **95**, 1595 – 1605.

Leeson, L. J. (1995). In Vitro/In Vivo Correlations. *Drug Information Journal* **29**, 903–915.

Lindsey, J. K., Jones, B., and Jarvis, P. (2001). Some statistical issues in modelling pharmacokinetic data. *Statistics in Medicine* **20**, 2775–2783.

Margolskee, A., Darwich, A. S., Galetin, A., Rostami-Hodjegan, A., and Aarons, L. (2016). Deconvolution and IVIVC: exploring the role of rate-limiting conditions. *The AAPS Journal* **18**, 321–332.

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (eds.), *Handbook of Markov Chain Monte Carlo*, chapter 5. CRC Press New York, NY.

OECD (2004a). Guidance document for the conduct of skin absorption studies number 28; OECD series on testing and assessment. *Adopted at 35th Joint Meeting August 2003.*

OECD (2004b). Guideline (428) for the testing of chemicals; Skin absorption: in vitro Method. *Adopted on 13th April 2004.* https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecdtg428-508.pdf.

O'Hara, T., Hayes, S., Davis, J., Devane, J., Smart, T., *et al.* (2001). In Vivo–In Vitro Correlation (IVIVC) Modeling Incorporating a Convolution Step. *Journal of Pharmacokinetics and Pharmacodynamics* **28**, 277–298.

Papaspiliopoulos, O., Roberts, G. O., and Sköld, M. (2007). A general framework for the parametrization of hierarchical models. *Statistical Science* **22**, 59–73.

Pinheiro, J. and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Statistics and Computing. Springer New York.

Piotrovskii, V. K. (1987). The use of Weibull distribution to describe the in vivo absorption kinetics. *Journal of Pharmacokinetics and Biopharmaceutics* **15**, 681–686.

Qiu, J., Martinez, M., and Tiwari, R. (2016). Evaluating In Vivo-In Vitro Correlation Using a Bayesian Approach. *The AAPS Journal* **18**, 619–634.

Rosenbaum, S. (2011). *Basic Pharmacokinetics and Pharmacodynamics: An Integrated Textbook and Computer Simulations*. Wiley, Hoboken.

Sheiner, L. B., Beal, S., Rosenberg, B., and Marathe, V. V. (1979). Forecasting individual pharmacokinetics. *Clinical Pharmacology & Therapeutics* **26**, 294–305.

SCCP (2006). Opinion on Basic criteria for the in vitro assessment of dermal absorption of cosmetic ingredients. *Adopted by the SCCP at 7th plenary meeting 28 March 2006.* https://goo.gl/hARU6a.

Stan Development Team (2018). *RStan: the R interface to Stan*. *R* package version 2.17.3. http://mc-stan.org/.

Tozer, T. N. and Rowland, M. (2006). *Introduction to Pharmacokinetics and Pharmacodynamics: The Quantitative Basis of Drug Therapy.* Lippincott William & Wilkins, Philadelphia. edition.

Yu, Z., Schwartz, J. B., Sugita, E. T., and Foehl, H. C. (1996). Five Modified Numerical Deconvolution Methods for Biopharmaceutics and Pharmacokinetics Studies. *Biopharmaceutics & Drug Disposition* **17**, 521–540.

Zhang, L., Beal, S. L., and Sheiner, L. B. (2003). Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *Journal of Pharmacokinetics and Pharmacodynamics* **30**, 387–404.

Zhou, H. (2003). Pharmacokinetic strategies in deciphering atypical drug absorption profiles. *The Journal of Clinical Pharmacology* **43**, 211–227.

