The AAPS Journal (2019) 21: 97
DOI: 10.1208/s12248-019-0366-2

Research Article

Evaluation of Assumptions Underpinning Pharmacometric Models

Qing-Xi Ooi,[1,3] Daniel F. B. Wright,[1] Geoffrey K. Isbister,[2] and Stephen B. Duffull[1]

Received 23 May 2019; accepted 9 July 2019; published online 5 August 2019

Abstract. Assumptions inherent to pharmacometric model development and use are not routinely acknowledged, described, or evaluated. The aim of this work is to present a framework for systematic evaluation of assumptions. To aid identification of assumptions, we categorise assumptions into two types: implicit and explicit assumptions. Implicit assumptions are inherent in a method or model and underpin its derivation and use. Explicit assumptions arise from heuristic principles and are typically defined by the user to enable the application of a method or model. A flowchart was developed for systematic evaluation of assumptions. For each assumption, the impact of assumption violation ('significant', 'insignificant', 'unknown') and the probability of assumption violation ('likely', 'unlikely', 'unknown') will be evaluated based on prior knowledge or the result of an additional bespoke study to arrive at a decision ('go', 'no-go') for both model building and model use. A table of assumptions with standardised headings has been proposed to facilitate the documentation of assumptions and evaluation of results. The utility of the proposed framework was illustrated using four assumptions underpinning a top-down model describing the warfarin-coagulation proteins' relationship. The next step of this work is to apply the framework to a series of other settings to fully assess its practicality and its value in identifying and making inferences from assumptions.

KEY WORDS: assumption; evaluation; assessment; pharmacometrics; models.

INTRODUCTION

All pharmacometric models are underpinned by assumptions. The validity of any inference drawn from a model depends on the appropriateness and likely impact of the underlying assumptions (1). This makes assumption evaluation an integral part of model building and model use.

The importance of assumption evaluation is well-recognised. Current guidelines by the Food and Drug Administration (FDA) (2), the European Medicines Agency (EMA) (3,4), and the European Federation of Pharmaceutical Industries and Associations (EFPIA) (1) stipulate that all assumptions inherent to model development and model application should be explicitly stated in the data analysis plan and study report. A transparent description of how these assumptions were assessed and what impact they might have on model inferences is also recommended. Here, the recommendations pertaining to acknowledgement, evaluation, and documentation of assumptions are applicable regardless of the type of analyses including pharmacokinetic (PK) analysis (2,3,5–8), pharmacokinetic and pharmacodynamic (PKPD) analysis (9,10), quantitative systems pharmacology (QSP) modelling (11), and physiologically based pharmacokinetic (PBPK) modelling (12–15).

The assumptions inherent to model development are not routinely acknowledged, evaluated, or documented in journal articles reporting pharmacometric analyses. This is also apparent in the analyses submitted for regulatory review, where the EMA outlined the lack of transparent description of influential assumptions and an ineffective evaluation or reporting of the impact of assumptions on model inference to be a major limitation (16). All these form an unequivocal barrier for effective model use and regulatory review.

While the importance of assumption evaluation is well-recognised, how these assumptions should be systematically approached and effectively assessed has received limited attention. Karlsson and colleagues used the development of a population PK model as an exemplar to demonstrate how assumptions linked to model building may be evaluated (7). The authors exhaustively stated the assumptions associated with the PK model, gave examples on how violations of assumptions can be detected, and introduced new models that allow some standard assumptions to be relaxed (7). In a white paper published by the EFPIA Model-Informed Drug

Electronic supplementary material The online version of this article (https://doi.org/10.1208/s12248-019-0366-2) contains supplementary material, which is available to authorized users.

[1] School of Pharmacy, University of Otago, 63 Hanover Street, Dunedin, 9016, New Zealand.
[2] School of Medicine and Public Health, University of Newcastle, Newcastle, NSW, Australia.
[3] To whom correspondence should be addressed. (e–mail: qingxiooi.pmx@gmail.com)

Discovery and Development (MID3) Workgroup (1), the importance of assumption evaluation and documentation was highlighted. Here, assumptions were broadly classified into five principal categories: (a) pharmacological assumptions, (b) physiological assumptions, (c) disease assumptions, (d) data assumptions, and (e) mathematical or statistical assumptions. Importantly, a table of assumptions was proposed for adoption by modellers to document assumptions, uncertainties, and impact of assumption violations (1).

The overarching goal of this work was to present a framework for the systematic evaluation of assumptions in order to encourage greater transparency in the description, evaluation, and documentation of influential assumptions for model building and model application. The specific objectives of this work were (a) to define the concept of an assumption, (b) to develop a flowchart for systematic evaluation of assumptions, (c) to propose a standard table for documentation of assumptions and evaluation results, and finally, (d) to illustrate the use of the proposed framework using a model building case example.

WORKFLOW

Existing frameworks (1,7) were used as a starting point for this work. The frameworks were expanded based on an evaluation of the risk management literature, expert opinion, and logical reasoning.

A simple three-step workflow is proposed for the evaluation of model-related assumptions: (a) step 1: identify assumptions (via the definition Table I and categorisation Table II); (b) step 2: evaluate assumptions and mitigate risks and uncertainties (via the flowchart Fig. 1); and (c) step 3: document assumptions, evaluations, and results (via the assumptions Table IV). Each step and the associated tools will be discussed in detail below.

STEP 1: IDENTIFY ASSUMPTIONS

Definition of an Assumption and Related Terms

For the purpose of this work, an assumption is defined as an unsubstantiated claim (*belief*) about a system that is required (*needed*) to be made in order for the system to be manipulated in a manner that is advantageous to the modeller (*model building or model use*). An assumption therefore is a belief that is needed to build or use a model.

A subset of these assumptions can be defined as hypotheses which, by definition, are testable. An assumption is distinguished from an axiom (a self-evident truth), a theorem (a proven belief [a truth]), and a limitation (a boundary beyond which the assumption no longer holds).

More elaborate definitions of these terms with examples are provided in Table I. Here, it is acknowledged that there may be some nuanced differences in the adopted definitions compared to that defined in other fields of study (e.g. mathematics, philosophy, psychology).

Categorisation of Assumptions

It is not the intention of this paper to present an exhaustive list of model-related assumptions. Indeed, the very existence of some assumptions may not be immediately obvious to modellers or may be so obvious to modellers that they become desensitised to the implications. This represents a major barrier to assumption evaluation and subsequent mitigation of risks and uncertainties.

Given these challenges, we propose a systematic approach that will allow modellers to identify assumptions based on how the assumption might arise. This requires that we distinguish two main categories of assumptions: 'implicit' and 'explicit' assumptions. A more detailed explanation and examples of implicit and explicit assumptions are provided below and in Table II. This classification is essential to help the modeller identify assumptions that may otherwise not be considered.

Implicit Assumptions

Implicit assumptions arise from an intrinsic component or aspect of a method or model. They are *not* defined by the user but are based on the science that underpins the method or model. Some examples might include the following: (a) the Cockcroft-Gault equation (17) carries the implicit assumption that the serum creatinine used in the equation is at steady-state; (b) the analysis of variance (ANOVA) requires the errors to be normally distributed. Here, a method or model may become invalid if the conditions required for the founding assumptions to be valid have not been met. Since the assumption attached to the method or model is not caused by the modeller, then it is not necessarily obvious, this type of assumption is termed *implicit*.

Explicit Assumptions

Explicit assumptions arise from the application of a method or model. These assumptions are defined by the user (hence are termed *explicit*) and may function as a heuristic solution to a problem although this often necessitates additional assumptions of appropriateness. Usually, explicit assumptions arise due to a need, lack of alternative methods or models, lack of knowledge of the true mechanism, and parsimony. Some examples might include (a) the recorded blood sampling time and dosing time are assumed to be accurate for model building and (b) the use of an E_{max} process in a pharmacometric model as a heuristic solution to describe an otherwise unknown mechanism of drug effect.

Explicit assumptions may include those that were originally implicit. This occurs when the knowledge of the (implicit) assumptions itself is matched to the structure of the problem to inform the choice of a method or a model. For example, since binding equilibrium is achieved quicker than changes in ligand concentration, application of an equilibrium binding model, e.g. an E_{max} model, may be considered appropriate. The implicit assumption associated with the equilibrium binding model, the ligand-receptor binding, is at equilibrium; the explicit assumption is that this is true for the system under study. Implicit and explicit assumptions are therefore not necessarily mutually exclusive.

Table I. Definition and examples of an assumption and related terms.

Term	Definition	Examples (not exhaustive)
Belief	A belief connotes a proposition that is considered to be true but for which substantive evidence is not available.	Example 1: Constant receptor system The turnover of receptors is negligible if the observation time window is relatively short in in-vitro experiments. Example 2: Metabolic pathways All metabolic pathways of a drug are known.
Assumption	An assumption is a belief where needs define its existence and necessitate its use.	Example 1: Accurate recorded sampling time The recorded sampling time is free of error Example 2: $\varepsilon \sim \text{iid } N(0, \sigma^2)$ All error terms ε are independent and identically distributed (iid) from a normal distribution with a mean of zero and a variance of σ^2 Example 3: GFR $\approx$ *e*GFR The estimated glomerular filtration rate (eGFR) obtained from the Cockcroft-Gault equation (17) provides an unbiased measure of the actual glomerular filtration rate (GFR) of a patient.
Limitation	A limitation represents the boundary beyond which the use of an assumption is unsound or invalid.	Example 1: Linear exposure-response model The assumed linear relationship between a drug's exposure and response may not hold beyond the dose range modelled. Example 2: Small sample size The sample size for the study is small thus potentially limiting generalisability of the study's results to the intended population. Example 3: Making inferences by extrapolation Inferences drawn from extrapolation about the clinical effects of dosing beyond the scope of the model may not be valid.
Hypothesis	An hypothesis is a testable assumption.	Example 1: Bioequivalence A given generic drug is bioequivalent to the original brand-name drug. Example 2: Hepatic clearance The fraction of a drug administered that is metabolised is 0.4. Example 3: Active compound Only the parent drug is pharmacologically active.
Axiom	An axiom is a belief that is regarded as self-evidently true.	Example 1: $CL \in \mathbb{R}^+$ Clearance of a drug (CL) is a positive quantity. Example 2: Volume of distribution Plasma volume represents the lower limit of the volume of distribution of any intravenously administered drugs.
Theorem	A theorem is a belief that is proven to be true based on logical or mathematical reasoning.	Example 1: Bayes' theorem Let X and Y be two events. Bayes' theorem states as follows: $P(Y\|X) = \frac{P(Y)P(X\|Y)}{P(X)}$. Example 2: Central limit theorem Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution (any distribution) with mean μ and finite variance σ^2. If the sample size n is sufficiently large, then the sampling distribution of the sample mean ($\overline{X}$) approaches a normal distribution with a mean of μ and a variance of $\frac{\sigma^2}{n}$.

STEP 2: EVALUATE ASSUMPTIONS AND MITIGATE RISKS AND UNCERTAINTIES (FLOWCHART)

The typical workflow for assumption evaluation was generalised into a qualitative flowchart (see Fig. 1). Assumptions, be it implicit or explicit, are entered into the flowchart one-at-a-time via the entry point labelled as "Start" in Fig. 1 for systematic evaluation of the impact of assumption violation (I) and the probability of assumption violation (P). The outcome of the evaluation using the flowchart is summarised using a decision matrix table (see Table III). Here, both I and P are evaluated for their influence on (a) an internal component of model building (termed *internal evaluation* and represented by subscript [i]) or (b) an external use of the model (termed *external evaluation* and represented by subscript [e]). The outcomes of the flowchart include the potential go-/no-go decision for internal model building or external use of the model, the acknowledgement of a limitation of the model or simply the acknowledgement of the assumption and continued model use. Internal and external evaluations should not be confused with implicit and explicit assumptions. Implicit and explicit assumptions

Table II. Definition and examples of implicit and explicit assumptions.

	Implicit assumption	Explicit assumption
Origin	Arise from an intrinsic component of a method or model.	Arise from heuristic principles or the application of a method or model.
Definition	1. The assumptions are inherent in a method or model and underpin its derivation and use. 2. The assumptions are *not* defined by the user but originate from the underpinning science.	1. The assumptions are required for the application of a method or model. 2. The assumptions are defined by the user and may function as a heuristic solution to a problem. Usually they arise due to a pragmatic need, lack of alternative methods or models, lack of knowledge of the true mechanism, or parsimony.
Observations	1. The assumptions are not always apparent and hence are termed *implicit*.	1. In many cases, the need for assumptions is based on a gap in knowledge, which requires an imputation. 2. In all cases, these assumptions are made by the user and hence are termed *explicit* (even if the user does not directly acknowledge the assumption).
Notes	All structural models (e.g. in pharmacokinetic-pharmacodynamic (PKPD) modelling) have some implicit assumptions. These are typically related to the model structure itself and also the alternative structures that are not permitted within the current framework. For instance, a two-compartment model with elimination occurring from the central compartment implicitly assumes that the eliminating organ(s) is located in the central compartment.	Implicit and explicit assumptions are *not* always mutually exclusive.
Assumption examples	Assumptions related to the derivation of an equation: 1. Cockcroft-Gault equation (17) carries the implicit assumption of steady-state serum creatinine. 2. The E_{max} model implicitly assumes that the total ligand concentration is in excess of the concentration of the receptors. Receptor binding therefore has a negligible effect on the free ligand concentration. Assumptions require for a method to work: 1. The analysis of variance (ANOVA) requires the errors in the predictions to be normally distributed. 2. The maximum likelihood (ML) estimation method typically requires the errors to be independent and identically distributed (iid). Assumptions of common pharmacometric models: 1. The turnover model typically assumes that the precursor pool for the physiological intermediate is abundant. 2. The relationship that specifies $CL = \frac{D \times F}{AUC}$ is predicated on elimination occurring from the central compartment. Note that AUC is the area under the concentration time curve, CL drug clearance, D dose, and F bioavailability. 3. Blood flow is constant and non-pulsatile	Heuristic assumptions that arise due to need, lack of alternative method or model, and parsimony: 1. Need/lack of alternative method or model: The recorded blood sampling time and dosing time are accurate. 2. Need/lack of alternative method or model: All metabolic pathways of a drug are known and only the parent drug is pharmacologically active. 3. Parsimony: For a particular PKPD dataset, a log-linear model provides an appropriate description of the relationship between drug concentration and effect. Explicit assumptions that are originally implicit: 1. That binding equilibrium is quicker than changes in concentration, hence application of an E_{max} model is appropriate. 2. Since the sampling of the drug concentrations is sparse, the error terms are unlikely to be autocorrelated, hence the use of the ML estimation method is appropriate. Assumptions of appropriateness: 1. The use of the Cockcroft and Gault equation to quantify the glomerular filtration rate is appropriate for patients with normal or chronic kidney disease and not acute kidney injury. 2. Application of an immediate effect model is acceptable for steady-state data even if the drug is known to have a delayed effect. 3. Application of a perfusion rate-limited, well-stirred, one-compartment tissue model in a PBPK model is appropriate for a drug that distributes freely and instantly across the membrane barriers
Limitation examples	Use of Janmahasatian's equation (18) to derive the fat free mass (FFM) for patient groups other than those of European descent. This assumption is implicit since it is defined in the original expression of the model.	Application of FFM equation to a patient who has recently lost weight via bariatric surgery. This assumption is explicit as the weight changes caused by surgery are designed to change body composition.

relate to the nature of the assumptions. We make the classification to help the modeller to identify assumptions that may otherwise not be considered. It is therefore not the intention here for modellers to categorise assumptions into these categories in order to work through their risk management strategy or as a separate means to evaluate assumptions.

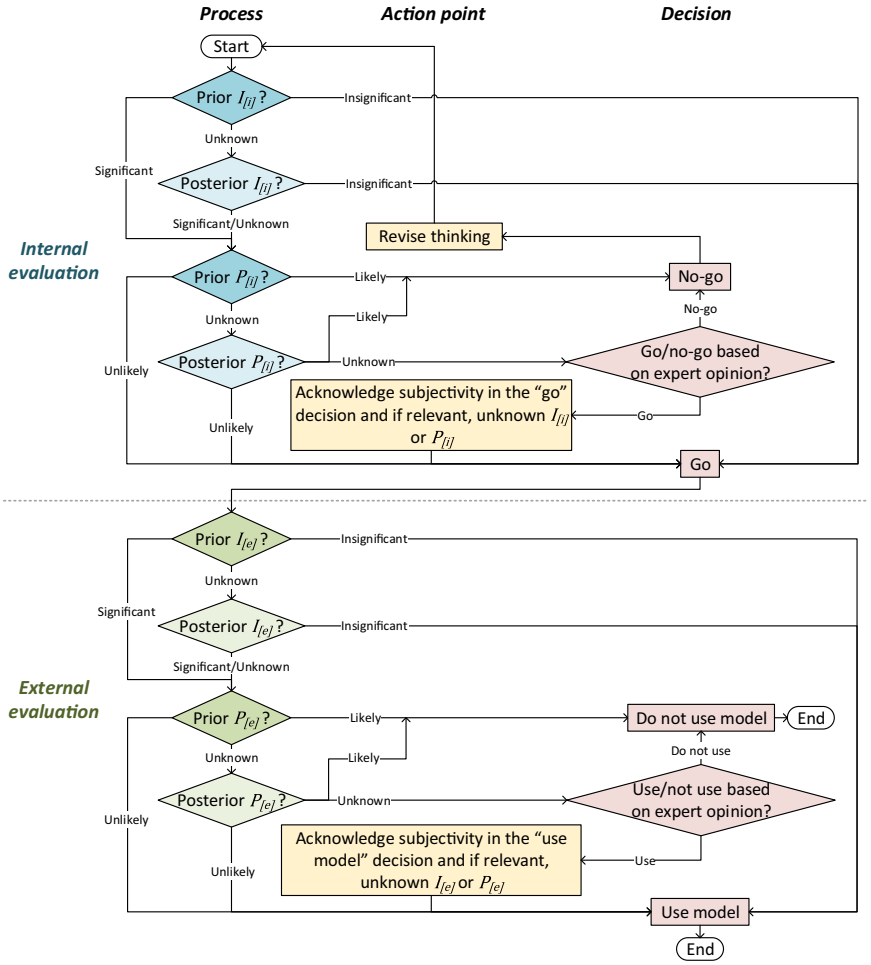

Fig. 1. Flowchart for systematic evaluation of model assumptions. *I* is the impact and *P* is the probability of assumption violation. The subscript [*i*] relates to internal evaluation (an evaluation relating to model development) and [*e*] external evaluation (an evaluation relating to use of the model after development)

Hence, we propose that assumptions, be it implicit or explicit, should be evaluated internally and/or externally using the flowchart during model development and application.

Descriptive ratings (significant, insignificant, likely, unlikely, unknown) are used for the purpose of this work. The choice of rating will depend on the relative magnitude of the risks and implications of go-/no-go which will need to be decided and documented on a case-by-case basis. Rating assignments can be based on either or both qualitative (e.g. logic, expert opinion, literature review) and quantitative (e.g. statistics, diagnostic plots) results. Here, a generation of qualitative evidence may potentially involve subjective interpretation of available resources and may therefore be dependent on the expertise, past experience, and resources of modellers or model users. On the other hand, quantitative evidence tends to be more objective. For instance, impact of assumption violation (*I*) may be assessed in a sensitivity analysis with respect to (a) model fits based on a change in the objective function value (OFV) (*I*= significant if $\Delta OFV \leq -3.84$) and (b) parameter estimates based on their confidence intervals (*I*= significant if the estimates of the same parameter has non-overlapping confidence interval). More examples of how ratings are assigned can be found in 'Application'.

The evaluation of *I* and *P* as well as the internal versus external evaluation are further discussed in the next following sections. The utility of the flowchart for decision-making is illustrated in 'Application'.

Evaluation of *I* and *P*

For each assumption, evaluation of the impact of assumption violation (*I*) as *significant*, *insignificant*, or *unknown* precedes the evaluation of the probability of assumption violation (*P*) as *likely*, *unlikely*, or *unknown*. If *I* is significant or unknown, then *P* is evaluated. In this hierarchical approach, *P* is otherwise not tested. Here, the ratings for *I* and *P* are rated based on either prior knowledge (e.g. existing literature, logical reasoning) or, in the absence of prior information, posterior knowledge which results from a bespoke study (e.g. experimental work, simulation study, sensitivity analysis).

In the special case when *I* is significant or unknown and *P* is unknown, it is recommended to make the go-/no-go decision based on the expert opinion. The expert opinion of an individual represents the collection of all unconscious-subconscious-conscious-learned experience. Here, it is

Table III. Decision matrix based on the impact and probability of assumption violation to evaluate if it is a go or no-go decision for model building (internal evaluation) and model use (external evaluation). Note here that when the impact of assumption violation is insignificant, the probability of assumption violation is not evaluated, thereby leading to an expedited go decision

		Probability of assumption violation (*P*)			
		Likely	Unlikely	Unknown	Not evaluated
Impact of assumption violation (*I*)	Significant	No-go	Go	Go/no-go based on expert opinion	NA
	Insignificant	Go	Go	Go	Go
	Unknown	No-go	Go	Go/no-go based on expert opinion	NA

important to recognise and acknowledge that the decision is based on opinion and should new information become available, the decision should be reviewed and revised if necessary. We therefore use the term expert opinion as a placeholder for decision-making, until a better source becomes available.

Internal Versus External Evaluation

Internal Evaluation

For *internal* evaluation, both impacts of the assumption violation (*I*) and probability of assumption violation (*P*) are evaluated with respect to the performance of the model. Here, the goal of the assessment of *I* is to quantify the influence of a violation in the assumption on the resulting model fit and evaluation of *P* is to assess the corresponding chance of assumption violation. The evaluation concludes with a go-/no-go decision.

External Evaluation

External evaluation is carried out in a similar manner except that it is with reference to an external aim of interest, which typically involves simulation from the final model and extrapolation beyond the original scope of the model. Evaluation concludes with a go-/no-go decision.

STEP 3: DOCUMENT ASSUMPTIONS, EVALUATIONS, AND RESULTS (TABLE OF ASSUMPTIONS)

A tabular approach to documentation is proposed (with particular emphasis on recording information of interest) to enhance the effectiveness of communicating assumptions. The table is divided up into three components, *Impact*, *Probability*, and *Decision*. The impact and probability of assumption violation components are further divided into *Methods*, *Results*, and *Rating*. The structure of the table of assumptions, its components, and a generic description of each component are shown in Table IV.

APPLICATION

The utility of the flowchart and table of assumptions was illustrated using a top-down case example (i.e. developing a model based on data and usually with parsimony in mind). In the case example, a joint model for six vitamin K-dependent coagulation proteins after warfarin dosing in 17 patients was developed (18). The model consists of a common one-compartment PK model with first-order absorption and elimination for warfarin that is linked to a parallel series of six inhibitory E_{max} models and turnover models for the coagulation proteins. The system of ordinary differential equations is shown in **Appendix** and the model schematic is shown in **Supplemental 1**.

Selected assumptions underpinning the joint model are shown here as examples to illustrate the use of the flowchart. Four application examples are included to illustrate internal evaluation (2 examples) and external evaluation (2 examples). The evaluation results of these assumptions are described in the next sections. A summary of the evaluation is also provided in Table V. Other assumptions underpinning the joint model were also evaluated. For brevity, these are not presented here but are summarised in **Supplemental 2**.

Internal Evaluation (Model Building)

Assumption Example 1

An assumption underpinning the joint model development was that the residual error terms, ε, are normally distributed with a mean of zero and a variance of σ^2. The modelling platform (NONMEM® Version 7.2, ICON Development Solutions, Ellicott City, MD, USA) is reported to be sensitive to a misspecification in the residual error model; therefore, parameter estimates, in particular, random-effect parameters, may be biased (19). For these reasons, the impact of assumption violation (*I*) was conservatively rated as significant.

The probability of assumption violation (*P*) was subsequently evaluated. Since no prior information was available, *P* was assessed by comparing the distribution of the conditional weighted residuals (CWRES) to a standard normal distribution using quantile-quantile (Q-Q) plots and one-sample Kolmogorov-Smirnov goodness-of-fit tests. Here, CWRES represent a composite measure of the additive and proportional error terms. In Fig. 2, the majority of the CWRES quantiles fall on the reference line that corresponds to a standard normal distribution. In addition, based on the Kolmogorov-Smirnov tests, CWRES for factors II ($T = 0.0315$, $p = 0.998$), IX ($T = 0.0895$, $p = 0.173$), X ($T = 0.0767$, $p = 0.328$), proteins C ($T = 0.0588$, $p = 0.660$), and S ($T = 0.0432$, $p = 0.933$) do not have a distribution that is significantly different from a standard normal distribution. An

Table IV. Table for documentation of assumption and their evaluation.

Assumption	Impact (*I*)			Probability (*P*)			Decision
	Methods	Results	Rating	Methods	Results	Rating	
State the assumption. Each row should contain only one assumption. Internal and external evaluations should be placed in separate rows.	State if *I* is not testable. If testable, state if the evaluation is based on prior knowledge or posterior test then outline the methods used briefly.	Summarise the evaluation results. Obtain supporting information for the rating assigned.	Rate *I* based on the evaluation results. The three possible ratings are *S* for significant, *NS* not significant, and *UK* unknown.	State if *P* is testable. If testable, state if the evaluation is based on prior knowledge or posterior test then outline the methods used briefly.	Summarise the evaluation results. Obtain supporting information for the rating assigned.	Rate *P* based on the evaluation results. The three possible ratings are *L* for likely, *UL* unlikely, and *UK* unknown.	State if it is a go or no-go decision for model building (internal evaluation) or model use (external evaluation).

exception was the CWRES for factor VII ($T = 0.111$, $p = 0.0458$), although visually the Q-Q plot does not look significantly worse than the other plots. Taken together, these results suggest that CWRES are likely to have a standard normal distribution. In other words, *P* of assumption violation was considered to be *unlikely*. *Evaluation result.* According to the flowchart (Fig. 1), significant *I* and low *P* give a go decision for model building.

Assumption Example 2

The data used to conduct the analysis of coagulation proteins for this illustrative example did not include a record of the warfarin dosing time. Therefore, a daily dose time of 6 p.m. was assumed for all patients based on the study protocol. In the absence of prior knowledge regarding the impact of assumption violation (*I*), *I* was evaluated a posteriori. A sensitivity analysis was conducted whereby different daily dose times of 8 a.m., 10 a.m., 12 p.m., 2 p.m., 4 p.m., 8 p.m., and 10 p.m. were used as alternative imputation values for dose time. The resulting model fits and parameter estimates were compared to the original model. Identical objective function values (OFV) and final parameter estimates were obtained thereby indicating insensitivity of model fits and parameter estimations to the different imputed daily dose time. In this case, *I* was considered insignificant and therefore unimportant, and as a result, the probability of assumption violation (*P*) was not evaluated. *Evaluation result.* According to the flowchart (Fig. 1), this leads to an expedited *go* decision for model building.

External Evaluation (Model Use)

For illustrative purposes, we will assume a hypothetical analysis goal to use the joint model for coagulation proteins to predict clinical scenarios beyond what is supported by the data. While a contentious issue methodologically, we may wish to explore whether the model is sufficiently mechanistic to predict new scenarios and to determine the limits of predictive performance.

Assumption Example 1

The joint model was developed based on data from patients who did not receive vitamin K doses or supplementation. The vitamin K intake in these patients is normal and assumed constant. The effect of warfarin on the vitamin K cycle was broadly represented by an inhibitory E_{max} model. Here, it is assumed that warfarin binding to the vitamin K epoxide reductase (VKOR) enzyme is reversible and that the binding affinity is time-invariant (i.e. the potency parameter, $IA_{50,\ B}$ is a constant).

In the current example, the hypothetical analysis goal is to use the joint model to predict the time course of coagulation protein concentrations in patients receiving vitamin K supplementation in addition to warfarin.

Based on logical reasoning, a deviation from the assumptions, for instance, if warfarin has a different binding affinity to VKOR and hence has a different (or variable) $IA_{50,}$

Table V. Four assumption examples and their evaluation.

Assumption	Impact (*I*)			Probability (*P*)			Decision
	Methods	Results	Rating	Methods	Results	Rating	
Example 1 for internal evaluation: For each coagulation protein, the errors, ε ($n \times 1$), are assumed normally distributed with a mean of zero and a variance of σ^2	Previous evaluation of impact of non-normality available in the literature (19).	Parameter estimation for maximum likelihood estimation is sensitive to specification of the residual error. Misspecification may result in biased parameter estimates and predictions.	S	The distribution of CWRES was compared to a standard normal distribution by using a Q-Q plot and a one-sample Kolmogorov-Smirnov test.	CWRES of all coagulation proteins do not have a distribution that is significantly different from a standard normal distribution.	U	Go for model building.
Example 2 for internal evaluation: The imputed daily dose time of 6 p.m. is appropriate and as a whole unbiased	A sensitivity analysis was conducted. Daily dose time of 8 a.m., 10 a.m., 12 p.m., 2 p.m., 4 p.m., 8 p.m., and 10 p.m. were considered.	The same model fits and parameter estimates were obtained regardless of the daily dose time used for imputation.	NS	NA	NA	NA	Go for model building.
Example 1 for external evaluation**:** Reversible binding and constant binding affinity *External population with vitamin K supplementation	Thought experiment and logical reasoning.	Two models that are exactly the same but differ in the value of potency parameter ($IA_{50,\ P}$) will give different predictions.	S	Prior knowledge	$IA_{50,\ P}$ is an apparent potency parameter that depends positively on the amount of vitamin K in the body. With vitamin K supplementation, $IA_{50,\ P}$ is almost definitely altered.	L	No-go for model use.
Example 2 for external evaluation: The dose-coagulation protein response relationship of warfarin is hyperbolic as described by the E_{max} model in the joint model. *Extrapolation beyond the original dose range used for model building	Thought experiment and logical reasoning.	Deviation from the stipulated E_{max} model is likely to result in discrepant predictions compared to the original model.	S	Prior knowledge	The inhibitory E_{max} models of the joint model are likely to be well-characterised as they were estimated based on a wide range of coagulation protein concentrations (i.e. from no inhibition to partial and almost complete inhibition of coagulation proteins production). The use of the model in high warfarin dose setting is justified.	U	Go for model use.

I is the impact of assumption violation, *L* means likely, *P* is the probability of assumption violation, *S* means significant, *NS* means not significant, *U* means unlikely, and *NA* not applicable

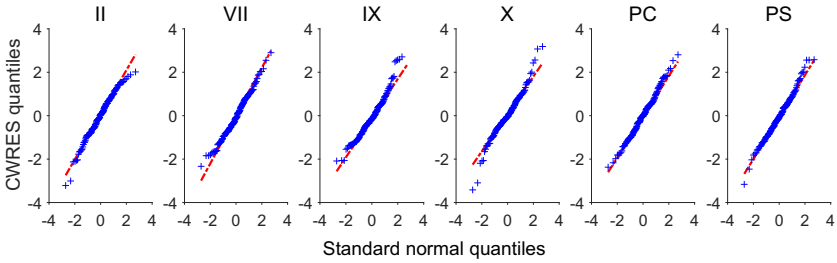

Fig. 2. The Q-Q plot for the CWRES of the joint model. CWRES is the conditional weighted residuals, PC is protein C, PS is protein S, and Q-Q plot refers to the quantile-quantile plot

$_P$, predictions from the model with the new $IA_{50,\ P}$ would be different from that produced by the original model. Impact of assumption violation (*I*) was rated as *significant*.

Next, the probability of assumption violation (*P*) was evaluated. $IA_{50,\ P}$ is an *apparent* potency parameter for warfarin that varies positively with the vitamin K level. Then, if the proposed cohort that the model is intended to predict into receives vitamin K supplements, the apparent $IA_{50,\ P}$ will increase. Due to the discrepancy in the apparent $IA_{50,\ P}$, simulations from the joint model will be biased as compared to the actual time course of coagulation protein concentrations in the new cohort. For these reasons, *P* was rated as *likely*. *Evaluation result.* According to the flowchart (Fig. 1), *I* = *significant* and *P* = *likely* result in a no-go decision for model use with respect to the hypothetical analysis goal. Indeed, further model development would be needed to achieve this aim.

Assumption Example 2

Warfarin inhibits VKOR that leads to a reduction in the production of functional coagulation proteins. In the joint model, the warfarin dose-coagulation protein response relationship was assumed to be hyperbolic and was described by an inhibitory E_{max} model. The parameter estimates relevant to the inhibitory E_{max} model are $IA_{50,II} = 1.83$ mg, $IA_{50,VII} = 1.86$ mg, $IA_{50,IX} = 2.41$ mg, $IA_{50,X} = 0.875$ mg, $IA_{50,PC} = 5.21$ mg, and $IA_{50,PS} = 3.35$ mg (18) (note: I_{max} was fixed to one).

In this example, the hypothetical analysis goal is to use the model to predict the time course of coagulation protein concentrations for warfarin doses higher than the original dose range used for model development.

Based on logical reasoning, if warfarin dose-coagulation protein response relationship deviates from the inhibitory E_{max} models estimated in the joint model, discrepancies in model predictions are expected. Impact of assumption violation (*I*) was rated as *significant*.

In the next step, the probability of assumption violation (*P*) was evaluated. Because the inhibitory E_{max} models in the joint model were estimated based on a wide range of coagulation proteins concentration, i.e. from no warfarin inhibition to partial and *almost complete* inhibition of coagulation proteins production, we believe that the inhibitory E_{max} models are sufficiently well-characterised. Hence, the predicted time course of coagulation protein concentrations, even for warfarin doses higher than the original dose range used for modelling, is likely to be reasonably accurate. *P* was rated as *unlikely*. *Evaluation result.* According to the flowchart (Fig. 1), significant *I* and low *P* give a go decision to use the model externally in the setting of a high-dose warfarin.

DISCUSSION

A framework for systematic evaluation of assumptions is proposed. The framework consists of a general workflow for assumption evaluation, definition, classification, and identification, as well as a flowchart for evaluating assumptions, and a proposed standardised table format for documentation. In this work, the utility of the framework is demonstrated using four examples of assumptions from a top-down example. Finally, this work illustrates how to determine go-/no-go decisions as well as providing criteria for determining when an assumption is deemed to be a limitation of the model.

Here, we propose categorising assumptions into implicit and explicit assumptions as a means to help modellers to identify assumptions systematically. Our categorisation is based on the origin of the assumption, that is how the assumption might arise. A different categorisation system has been suggested previously. In the white paper published by Marshall *et al.* (1), assumptions are categorised into six categories according to the nature of assumptions: (a) biological or physiological assumptions, (b) pathophysiological assumptions, (c) pharmacological or pharmaceutical assumptions, (d) experimental assumptions, (e) study conduct assumptions, and (f) statistical or mathematical assumptions. The definition and examples of these assumptions are summarised in **Supplemental 3**. We believe that these two classification systems of assumptions complement one another in the identification of assumptions and are therefore important *aide-mémoire* to modellers.

In this framework, we interrogate assumptions on the basis of their impact if violated and the probability of assumption violation. This is similar to the standard risk assessment matrix used in project management (20,21). We also simplify the problem by considering a linear workflow which obviates the need to consider all situations for all assumptions, for example, if an assumption has a low impact of assumption violation then we propose the modeller does not consider the probability. This contrasts with risk assessment in project management. We do this because biology is complicated, data is fraught, and models are simple, and hence the scale of the assumptions inherent in model development may be of high dimensions. In many cases, due to the low impact of many assumptions, it may be sufficient to determine that the

assumption has been recognised. Finally, we formalise the term *limitation* as it pertains to the application of a model. This term is now recommended to be reserved for situations in which assumption violation is associated with a significant impact and a high probability of assumption violation, and hence it is a limitation of the model that it cannot be used for that setting.

There have been few published works that address how model-related assumptions should be systematically approached and be effectively assessed. To date, only two published works are available that provide specific frameworks for evaluating model assumptions. The first work by Karlsson and colleagues used the development of a population PK model for moxonidine as an exemplar to demonstrate how assumptions intrinsic to model building may be evaluated (7). The second article by the EFPIA MID3 Workgroup (1) provides recommendations on how assumptions may be evaluated and documented. Noteworthy, an assumptions table was proposed for modellers to document assumptions, uncertainties, and impact of assumption violations. Our work is viewed as an extension of these existing frameworks. Interested readers are referred to these works for more examples and guidance on assumption evaluation. In this work, a one-stop framework that covers the typical workflow for assumption evaluation is described and includes identification, evaluation, and documentation of assumptions. With this, modellers are encouraged to describe, evaluate, and document influential assumptions for effective model building and model use.

There are a number of limitations to this work. The term *assumption* and related terms are defined within the context of model development and use. These are rather heuristic definitions employed to empower modellers to differentiate these closely related but potentially confusing terms. It is acknowledged that there may be nuanced differences in the adopted definitions compared to that defined in other fields of study (e.g. mathematics, philosophy, psychology). Although this work aims to assist modellers to evaluate assumptions objectively, the use and application of the framework for assumption evaluation is somewhat subjective. For instance, the choice of evaluation methods, the rigour in which assumptions are interrogated, and the rating assignments for the impact and probability of assumption violation are all subjective and dependent on the modellers' expertise, past experience, preference, available resources, and time constraints. In addition, the required first step for assumption evaluation is successful identification of pertinent assumptions. In this work, it is suggested for modellers to list the assumptions methodically according to their nature. However, despite the systematic approach suggested, identification of an exhaustive list of model-related assumptions is challenging. Due to the implicit nature of some assumptions and model complexity, it is possible for some assumptions to have gone unnoticed, i.e. unknown unknowns. Here, it remains unanswered as to how to mitigate the risk associated with these unknown unknowns.

This framework has been applied to a top-down approach. Although the examples are sufficient to illustrate how the proposed framework may be used for assumption evaluation, it is not at this point a form of validation. The robustness of the framework to the different modelling processes and model use settings as well as the utility and practicality of the framework require further exploration. An important next step of this work is therefore to apply the framework to other settings including that of bottom-up works (e.g. QSP model building).

CONCLUSION

A framework for systematic evaluation of assumptions is proposed and its utility is demonstrated using a model building example. The next step of this work is to apply the framework to a series of other settings to fully assess its practicality and its value in identifying and making an inference from assumptions. Finally, a simple electronic version, perhaps a web tool, would be useful in streamlining the next step.

APPENDIX

Assumptions underpinning a joint model for six vitamin K-dependent coagulation proteins (18) were evaluated. The joint model is described by the following system of ordinary differential equations:

Pharmacokinetic model:

$$\frac{dD}{dt} = -k_a \cdot D; \quad \begin{cases} t < \mathrm{LAG}, D_{t=0} = 0 \\ t \geq \mathrm{LAG}, D_{t=0} = \mathrm{Dose} \end{cases}$$

$$\frac{dA}{dt} = k_a \cdot D - \frac{\mathrm{CL}}{V} \cdot A; \quad A_{t=0} = 0$$

Pharmacodynamic model:

$$\frac{d\mathrm{II}}{dt} = R_{\mathrm{in,II}} \cdot I_{II} - k_{\mathrm{out,II}} \cdot \mathrm{II}; \quad \mathrm{II}_{t=0} = \frac{R_{\mathrm{in,II}}}{k_{\mathrm{out,II}}}$$

$$\frac{d\mathrm{VII}}{dt} = R_{\mathrm{in,VII}} \cdot I_{\mathrm{VII}} - k_{\mathrm{out,VII}} \cdot \mathrm{VII}; \quad \mathrm{VII}_{t=0} = \frac{R_{\mathrm{in,VII}}}{k_{\mathrm{out,VII}}}$$

$$\frac{d\mathrm{IX}}{dt} = R_{\mathrm{in,IX}} \cdot I_{\mathrm{IX}} - k_{\mathrm{out,IX}} \cdot \mathrm{IX}; \quad \mathrm{IX}_{t=0} = \frac{R_{\mathrm{in,IX}}}{k_{\mathrm{out,IX}}}$$

$$\frac{dX}{dt} = R_{\mathrm{in},X} \cdot I_X - k_{\mathrm{out},X} \cdot X; \quad X_{t=0} = \frac{R_{\mathrm{in},X}}{k_{\mathrm{out},X}}$$

$$\frac{d\mathrm{PC}}{dt} = R_{\mathrm{in,PC}} \cdot I_{\mathrm{PC}} - k_{\mathrm{out,PC}} \cdot \mathrm{PC}; \quad \mathrm{PC}_{t=0} = \frac{R_{\mathrm{in,PC}}}{k_{\mathrm{out,PC}}}$$

$$\frac{d\mathrm{PS}}{dt} = R_{\mathrm{in,PS}} \cdot I_{\mathrm{PS}} - k_{\mathrm{out,PS}} \cdot \mathrm{PS}; \quad \mathrm{PS}_{t=0} = \frac{R_{\mathrm{in,PS}}}{k_{\mathrm{out,PS}}}$$

$$P = \{\mathrm{II}, \mathrm{VII}, \mathrm{IX}, \mathrm{X}, \mathrm{PC}, \mathrm{PS}\}$$

$$I_P(t) = 1 - \frac{I_{\max,P} \cdot A(t)}{IA_{50,P} + A(t)}$$

$$R_{\mathrm{in},P} = P_{t=0} \cdot k_{\mathrm{out},P}.$$

Here, LAG, general lag parameter; CL , clearance of warfarin; D, warfarin dose; II, factor II; $IA_{50,P}$, warfarin amount in the body that gives half the maximum inhibitory effect; $I_{\max,P}$, maximum inhibitory effect; IX, factor IX; k_a,

first-order absorption rate constant of warfarin; $k_{out,P}$, first-order coagulation protein degradation rate constant; PC, protein C; PS, protein S; $R_{in,P}$, zero-order functional coagulation protein production rate; t, time; V, volume of distribution of warfarin; VII, factor VII; X, factor X; $P_{t=0}$, coagulation protein concentration at baseline.

REFERENCES

1. Marshall SF, Burghaus R, Cosson V, Cheung SY, Chenel M, DellaPasqua O, et al. Good practices in model-informed drug discovery and development: practice, application, and documentation. CPT Pharmacometrics Syst Pharmacol. 2016;5(3):93–122.
2. Food and Drug Administration (FDA). Guidance for industry: population pharmacokinetics 1999 [Available from: https://www.fda.gov/downloads/drugs/guidances/UCM072137.pdf. Accessed 15 April 2019.
3. European Medicines Agency (EMA). Guideline on reporting the results of population pharmacokinetic analyses 2007 [Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003067.pdf. Accessed 15 April 2019.
4. Manolis E, Brogren J, Cole S, Hay JL, Nordmark A, Karlsson KE, et al. Commentary on the MID3 good practices paper. CPT Pharmacometrics Syst Pharmacol. 2017;6(7):416–7.
5. Byon W, Smith MK, Chan P, Tortorici MA, Riley S, Dai H, et al. Establishing best practices and guidance in population modeling: an experience with an internal population pharmacokinetic analysis guidance. CPT Pharmacometrics Syst Pharmacol. 2013;2:e51.
6. Dykstra K, Mehrotra N, Tornøe CW, Kastrissios H, Patel B, Al-Huniti N, et al. Reporting guidelines for population pharmacokinetic analyses. J Pharmacokinet Pharmacodyn. 2015;42(3):301–14.
7. Karlsson MO, Jonsson EN, Wiltse CG, Wade JR. Assumption testing in population pharmacokinetic models: illustrated with an analysis of moxonidine data from congestive heart failure patients. J Pharmacokinet Biopharm. 1998;26(2):207–46.
8. Wade JR, Edholm M, Salmonson T. A guide for reporting the results of population pharmacokinetic analyses: a Swedish perspective. AAPS J. 2005;7(2):45.
9. Bonate PL, Strougo A, Desai A, Roy M, Yassen A, van der Walt JS, et al. Guidelines for the quality control of population pharmacokinetic-pharmacodynamic analyses: an industry perspective. AAPS J. 2012;14(4):749–58.
10. Jamsen KM, McLeay SC, Barras MA, Green B. Reporting a population pharmacokinetic-pharmacodynamic study: a journal's perspective. Clin Pharmacokinet. 2014;53(2):111–22.
11. Timmis J, Alden K, Andrews P, Clark E, Nellis A, Naylor B, et al. Building confidence in quantitative systems pharmacology models: an engineer's guide to exploring the rationale in model design and development. CPT Pharmacometrics Syst Pharmacol. 2017;6(3):156–67.
12. Jones H, Rowland-Yeo K. Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. CPT Pharmacometrics Syst Pharmacol. 2013;2:e63.
13. Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, et al. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. Clin Pharmacol Ther. 2015;97(3):247–62.
14. Rowland M, Peck C, Tucker G. Physiologically-based pharmacokinetics in drug development and regulatory science. Annu Rev Pharmacol Toxicol. 2011;51:45–73.
15. Zhao P, Rowland M, Huang SM. Best practice in the use of physiologically based pharmacokinetic modeling and simulation to address clinical pharmacology regulatory questions. Clin Pharmacol Ther. 2012;92(1):17–20.
16. (EMA) EMA. EFPIA-EMA modelling and simulation workshop (2011) report. London, United Kingdom; 2012.
17. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31–41.
18. Ooi QX, Wright DFB, Tait RC, Isbister GK, Duffull SB. A joint model for vitamin K-dependent clotting factors and anticoagulation proteins. Clin Pharmacokinet. 2017;56(12):1555–66.
19. Karlsson MO, Beal SL, Sheiner LB. Three new residual error models for population PK/PD analyses. J Pharmacokinet Biopharm. 1995;23(6):651–72.
20. Hubbard DW. The failure of risk management: why it's broken and how to fix it: John Wiley & Sons; 2009.
21. Woodruff JM. Consequence and likelihood in risk estimation: a matter of balance in UK health and safety risk assessment practice. Saf Sci. 2005;43(5):345–53.