Check for updates

Received: 17 May 2021 | Revised: 17 May 2021 | Accepted: 2 July 2021

DOI: 10.1002/psp4.12696

TUTORIAL

Pharmacometrics meets statistics—A synergy for modern drug development

Yevgen Ryeznik[1] | **Oleksandr Sverdlov**[2] | **Elin M. Svensson**[3,4] | **Grace Montepiedra**[5] | **Andrew C. Hooker**[3] | **Weng Kee Wong**[6]

[1]BioPharma Early Biometrics and Statistical Innovation, Data Science & AI, R&D Biopharmaceuticals, AstraZeneca, Gothenburg, Sweden

[2]Early Development Analytics, Novartis Pharmaceuticals Corporation, East Hanover, New Jersey, USA

[3]Department of Pharmacy, Uppsala University, Uppsala, Sweden

[4]Department of Pharmacy, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

[5]Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

[6]Department of Biostatistics, University of California Los Angeles, Los Angeles, California, USA

Correspondence
Yevgen Ryeznik, BioPharma Early Biometrics and Statistical Innovation, Data Science & AI, R&D Biopharmaceuticals, AstraZeneca, Gothenburg, Sweden.
Email: yevgen.ryeznik@astrazeneca.com

Funding information
G.M.'s work was supported by the Statistical and Data Management Center of the International Maternal Pediatric Adolescent AIDS Clinical Trials Network under the National Institute of Allergy and Infectious Diseases grant No. UM1 AI068616.

Abstract

Modern drug development problems are very complex and require integration of various scientific fields. Traditionally, statistical methods have been the primary tool for design and analysis of clinical trials. Increasingly, pharmacometric approaches using physiology-based drug and disease models are applied in this context. In this paper, we show that statistics and pharmacometrics have more in common than what keeps them apart, and collectively, the synergy from these two quantitative disciplines can provide greater advances in clinical research and development, resulting in novel and more effective medicines to patients with medical need.

> We build too many walls and not enough bridges.
>
> Isaac Newton (1642–1727).

INTRODUCTION

Modern drug development is a long, complex, and expensive enterprise.[1] In clinical drug development, randomized controlled trials (RCTs) have been the established as the "gold standard" of experiments to obtain cause–effect evidence that an investigational treatment is better than a standard of care. Statistical methodologies have been the primary engine to design and analyze data for clinical trials. As such, statisticians have traditionally taken the leadership position in providing scientific and technical expertise in data analytic tasks within clinical development programs. There is even an explicit regulatory requirement that "...the actual responsibility for all statistical work associated with clinical trials will lie with an appropriately qualified and experienced statistician...," who "...should have a combination of education/training and experience sufficient to implement the principles articulated in this (ICH E9) guidance."[2]

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. CPT: *Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics

However, many problems in drug development cannot be adequately solved using traditional statistical methods, and they require more elaborate approaches using physiology-based drug and disease models. Pharmacometrics (PMx) is a relatively new quantitative discipline that integrates drug, disease, and trial information to facilitate efficient implementation of drug development and/or regulatory decisions.[3] PMx provide tools for modeling and simulation of clinical trials and, because PMx incorporates biologically based mathematical modeling within a statistical framework, the synergy between statisticians and pharmacometricians can help solving complex problems in drug development.

Lewis Sheiner championed the "learn and confirm" paradigm for drug development.[4] He argued that the research questions should determine the appropriate analytic methods. In the "learn" phase, the goal is exploration and the focus should primarily be on estimation of dose- and exposure-response relationships, whereas the "confirm" phase should focus on testing relevant clinical research hypotheses. Within the clinical research and development (R&D) workflow, Sheiner distinguished two "learn–confirm" cycles: phase I to phase IIa (proof-of-concept), and phase IIb (dose ranging) to phase III (confirmatory studies). Furthermore, Sheiner advocated a continuous, model-based approach that integrates relevant accumulated knowledge to optimize decision making within clinical development programs. This is now known as model-informed drug development (MIDD).[5–7] Other terms, such as model-aided drug development (MADD) or model-based drug development (MBDD), have also been used in the literature.

It seems natural that PMx scientists and statistical scientists should strive to provide collaborative support to drug development teams. However, this is often not the case.[8,9] One delicate issue is which of the two functions—statistics or PMx—should take the leading role and coordinate the activities, and how the actual data analytic tasks should be handled. A tension between statisticians and pharmacometricians is documented in the following vivid passage that reflected on Stephen Senn's conversation with Lewis Sheiner during the European Cooperation in the Field of Scientific and Technical Research (COST) meeting in Geneva in 1997[8]:

> ...Lewis had criticized a particular dogma in drug development as being excessively negative, conservative, and, in consequence, inimical to the conduct of proper science. The battle lines were clear. On the one side were the forces of light: those who liked models that used biological insights, generally welcomed data from disparate sources (even those that had not arisen from tightly controlled randomized experiments), and were not afraid to try various bold and ingenious strategies for putting models and data together. On the other side were the forces of darkness: a bunch of dice throwers and hypothesis testers with an inane obsession with intention to treat...

More recently, there has been substantial progress in promoting the broad use of MIDD. PMx is an integral part of quantitative science groups in many biopharmaceutical companies. PMx scientists and statistical scientists are increasingly working together. The integration of fields in this context is viewed as an important and necessary attribute of modern drug development.[10]

In the current paper, we provide an overview and compare statistical and pharmacometric approaches to solving important problems in clinical drug development. Through examples, we show that each approach has its own strengths and limitations, and these differences provide unique opportunities for collaboration and synergy: for statisticians to learn more about assumption-rich models, and for PMx scientists to see application of these models over the broader spectrum of clinical trial methodologies. Our purpose is to highlight the interdisciplinary nature of modern drug development, the necessity for integration of different fields, and the importance of cross-functional collaboration. The target audience for this paper includes statisticians, pharmacometricians, and other quantitative scientists working in the biopharmaceutical industry, academia, and health authorities.

STATISTICS AND PHARMACOMETRICS: TWO WORLDS APART?

In what follows, we briefly discuss some major sources of difference between statistics and PMx —namely, the approaches to modeling and the types of problems in drug development each discipline is best suited to solve. We also highlight some areas where there is a good common ground between the disciplines. For a more in-depth discussion of the philosophical gaps and historical sources of tension between statisticians and pharmacometricians, see Senn[8] and Kowalski[9].

Statistical approach

The model building process is a fundamental step in statistical problem solving.[11] The choice of a model depends on the nature of the problem and the research question; for example, if we are interested in examining whether

the treated group responds better than the untreated group, controlling for other factors in a trial, the logistic or the probit model comes to mind. However, one should always seek the right balance between model complexity and goodness-of-fit to data. Albert Einstein's famous phrase "As simple as possible, but not simpler" is a good epitome of the principle of parsimony. As an illustration, the logistic model is typically the preferred choice because it is easier to both compute and interpret than the probit model.

The RCT is the backbone of evidence-based clinical drug development. Traditionally, RCTs are designed and analyzed using frequentist methods; however, a Bayesian paradigm is becoming increasingly more common,[12] especially in precision medicine development.[13] For an RCT designed using a frequentist paradigm, the following empirical model is widely known and commonly used. Suppose we want to compare the effects of two treatments, experimental (E) and control (C), with respect to some quantitative clinical outcome (Y). Eligible subjects are enrolled into the study sequentially and each subject is randomized into one of the groups, receiving either E or C. For the i^{th} subject, let $\delta_i = 1$ (or $\delta_i = 0$), if the treatment assignment is E (or C), with $i = 1, \cdots, n$. Assuming outcomes are normally distributed with group means μ_E and μ_C and common variance σ^2, a statistical model for the outcomes Y_i (conditional on δ_i's) is:

$$Y_i = \delta_i \mu_E + \left(1 - \delta_i\right) \mu_C + \varepsilon_i, \quad i = 1, \ldots, n \qquad (1)$$

assuming the random errors ε_i's are independent and identically distributed (i.i.d.) $N(0, \sigma^2)$. To compare the two groups E and C, one tests the hypothesis $H_0: \mu_E = \mu_C$ (i.e., treatment effects are the same) versus $H_1: \mu_E > \mu_C$ (E is better than C). A two-sample t-test statistic is:

$$T = \frac{\overline{Y}_E - \overline{Y}_C}{\sqrt{S_p^2 \left(\frac{1}{n_E} + \frac{1}{n_C}\right)}},$$

where $\overline{Y}_E$ and $\overline{Y}_C$ are the group sample means, S_p^2 is the pooled sample variance, and $n_E = \sum_{i=1}^{n} \delta_i$ and $n_C = n - n_E$ are the group sizes. The null hypothesis H_0 is rejected at significance level α if $T > t_{1-\alpha, n-2}$, where $t_{1-\alpha,n-2}$ is the $100\,(1-\alpha)^{th}$ percentile of the t-distribution with $n-2$ degrees of freedom. Equivalently, H_0 is rejected if the one-sided p value of the test is less than α.

At the study planning stage, the sample size n is chosen to ensure sufficiently high statistical power (typically ≥80%) to reject H_0 when the true mean difference $\mu_E - \mu_C$ is equal to some prespecified clinically meaningful value $\Delta > 0$. Equal (1:1) allocation is usually implemented, that is, $n/2$ patients are allocated to each treatment group.

The described working model is very simple, but it is still widely considered by clinical trial statisticians. It may be very appropriate for a phase III confirmatory clinical trial, where the goal is to obtain a definitive assessment of the treatment effect via the predefined hypothesis test.[2]

Two important remarks should be made here for RCTs. First, in practice, many analyses for RCTs may utilize a model for the primary outcome that is more elaborate than Equation (1) and use a technique other than a two-sample t-test. For example, Equation (1) can be extended to include clinically important baseline covariates and evaluated through an analysis of covariance (ANCOVA). In trials with longitudinal continuous outcomes, a linear mixed model with repeated measurements (MMRMs) and proper covariate adjustments is frequently used.[14] Trials with binary outcomes may utilize a generalized linear model (such as a logistic regression model), and trials with time-to-event outcomes could use a semiparametric model (such as Cox's proportional hazards model) for a prespecified primary analysis. Most of these approaches are based on the (generalized) linear model theory, and the data are analyzed using standard statistical inference techniques in which statisticians are well trained. In many circumstances, this should be sufficient to answer a primary research question.

Second, a major strength of the RCT is the use of randomization as a mechanism to allocate trial participants to treatment groups. Randomization mitigates various experimental biases, promotes comparability of treatment groups, and validates the use of statistical methods for inferential analysis of trial results.[15] In most cases, statistical inference following the RCT is carried out *conditionally* on the design (i.e., randomization is not accounted for in the analysis). In this case, one makes a major assumption of the invoked population model, namely, that within each treatment group the observed responses are independent and identically distributed.[16] An alternative analytical approach uses the randomization model, in which the experimental randomization itself forms the basis for statistical inference. This idea can be traced back to R. A. Fisher's work on the design of experiments.[17] Randomization-based tests yield results similar to likelihood-based analysis methods when the population model assumptions are satisfied, but the former tests may be more robust when the model assumptions are violated.[18]

Each of the described statistical approaches has merits and limitations. Some may be appropriate for testing a carefully defined clinical research hypothesis using the above framework for a phase III pivotal trial, but may be unsuitable for a variety of other studies throughout the drug development lifecycle that are driven by different experimental goals.

Pharmacometric approach

With a PMx approach, one starts a model building process with an understanding of the pharmacological and biological mechanisms underlying the phenomena of interest. Often times, the mechanisms are nonlinear. To fix ideas, let us consider a basic problem of modeling the drug pharmacokinetic-pharmacodynamic (PK/PD) relationship. For the PK part, we consider a one-compartment model with first order absorption and elimination described by the following differential equations[19]:

$$\begin{cases} \dfrac{dX_c(t)}{dt} = K_a X_a(t) - K_e X_c(t), \; X_c(0) = 0; \\ \dfrac{dX_a(t)}{dt} = -K_a X_a(t), \qquad X_a(0) = d. \end{cases} \tag{2}$$

In Equation (2), $X_c(t)$ is the amount of drug in the central compartment (e.g., blood) at time t after administration, $X_a(t)$ is the amount of drug at the absorption site at time t, d is the administered dose, and K_a and K_e are, respectively, the absorption and the elimination rate constants. PMx scientists frequently consider the relationship $K_e = \mathrm{CL}/V$, where V stands for the volume of distribution and CL (clearance) measures the volume of blood which is cleared for drug per time.

Let $C(t) = X_c(t)/V$ denote the drug concentration in the central compartment. A direct calculation shows the solution to Equation (2):

$$C(t) = \frac{K_a d}{K_a V - \mathrm{CL}} \left(e^{-\frac{\mathrm{CL}}{V} t} - e^{-K_a t} \right), \quad K_a > \frac{\mathrm{CL}}{V}, \quad t > 0 \tag{3}$$

In Equation (3), the unit for $C(t)$ is mass/volume, and the relationship has a meaningful biological interpretation. The parameters $\boldsymbol{\beta} = \left(\mathrm{CL}, V, K_a\right)'$ determine the drug PK theoretical process over time.

For the PD part, we consider the frequently used sigmoidal maximum effect ($E_{\max}$) model[20] for a continuous response, which can be some relevant biomarker related to the drug target engagement or mechanism of action:

$$R(t) = E_0 + \frac{E_{max} C(t)^h}{EC_{50}^h + C(t)^h}. \tag{4}$$

Here, E_0 is the effect when the drug concentration is zero, $E_{\max}$ is the maximum change in effect attributable to the drug, EC_{50} is the level of exposure to produce 50% of $E_{\max}$, and h is the slope coefficient (Hill factor) that determines the steepness of the sigmoidal curve.

Equations (3) and (4) together define a PK/PD relationship that links the effect of dose (d) on drug concentration $C(t)$ and on drug response $R(t)$ over time. In practice, an investigator may wish to study the process in a population of subjects, accounting for biological and other clinically important sources of variation among individuals. In a clinical research setting, this leads to the introduction of a nonlinear mixed effects model (NLMEM).

Suppose there are n subjects in the study. For the i^{th} subject, the dose d_i is administered at time zero and, subsequently, the drug concentrations are measured repeatedly at times $t_{i1}, \cdots, t_{i,m_i}$, $i = 1, \cdots, n$. Let C_{ij} denote the drug concentration of the i^{th} subject at time t_{ij}. Then one can postulate the following model:

$$C_{ij} = C(t_{ij}, \boldsymbol{\beta}_i) + \varepsilon_{ij}, \tag{5}$$

where $\boldsymbol{\beta}_i = (\mathrm{CL}_i, V_i, K_{ai})'$ are the parameters of the PK process $C\left(t_{ij}, \boldsymbol{\beta}_i\right)$ specific to the ith individual and ε_{ij} are independent measurement errors with mean zero and variances that may vary across i and j. Note that $\boldsymbol{\beta}_i$'s are assumed to vary across individuals in the population, and this variation can be decomposed into a systematic part or fixed effects determined by typical (population mean) values $\boldsymbol{\beta} = \left(\mathrm{CL}, V, K_a\right)'$ and a random effect part $\boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i3})'$, associated with the individual variability. One can model the components of $\boldsymbol{\beta}_i = \left(\mathrm{CL}_i, V_i, K_{ai}\right)'$ as $\mathrm{CL}_i = e^{b_{i1}} \mathrm{CL}$, $V_i = e^{b_{i2}} V$, $K_{ai} = e^{b_{i3}} K_a$, where $\boldsymbol{b}_i = \left(b_{i1}, b_{i2}, b_{i3}\right) \sim N(0, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega}$ is a 3×3 unstructured covariance matrix. With these assumptions, CL_i, V_i, and K_{ai} are lognormal random variables. There may also be effects associated with some important covariates, but we do not consider them in the model for the sake of simplicity.

To model individual PD responses, $C(t)$ in Equation (4) is replaced by C_{ij} from Equation (5) with additional subject random effects $\boldsymbol{\nu}_i = (E_{0i}, E_{max,i}, \mathrm{EC}_{50,i}, h_i)'$ and measurement errors. The resulting NLMEM is then fitted using a software such as NONMEM[21,22] to produce estimates of the parameters with the corresponding uncertainties.

The described PK/PD modeling approach has several advantages. It incorporates scientific knowledge through a mathematical model, and it explicitly captures individual PK/PD behavior via parameters $\boldsymbol{\beta}_i$ and $\boldsymbol{\nu}_i$, thereby allowing modeling of individual subject profiles with an assessment of the corresponding uncertainty. Obviously, the assumed model should be plausible.

Unlike the statistical approach ("Statistical Approach" section), which was described in the context of a phase III RCT comparing the treatment effects via hypothesis testing, the PMx approach described here (Pharmacometric Approach section) addresses a different research question, namely estimation of a PK/PD profile of the drug and potentially prediction of the drug effect over time for study participants. Importantly, the PD response variable must be chosen judiciously; often it would be some

short-term marker predictive of long-term clinical outcomes. It stands to reason that the PK/PD approach adds great value during the exploratory (phases I and II) parts of drug development; however, it can be also useful in the analysis of data from confirmatory phase III trials (e.g., by directly correlating drug exposure and long-term clinical outcomes of interest).

Convergence of pharmacometric and statistical approaches

There are good examples in drug development where statistical and PMx approaches are in concordance. Two such examples are dose–response (DR) and exposure–response (ER) modeling.

DR studies are typically conducted during phase II of clinical development, where the goals are to characterize DR relationships over a given range of doses and to identify dose(s) suitable for subsequent testing in phase III confirmatory trials. A conventional phase II DR study is a randomized, placebo-controlled, parallel group design evaluating the effects of prespecified dose levels of a drug $d_1 < d_2 < \cdots < d_K$ and the placebo ($d_0 = 0$) on a response (clinical outcome or biomarker). With a standard statistical approach, one would consider a single continuous response per patient, often selected at some predefined time point after a dose of the drug. An analysis of variance (ANOVA) model is then postulated as $Y_{ik} = \mu_k + \varepsilon_{ik}$, where μ_k is the mean response in group $k = 0, 1, \cdots, K$ and ε_{ik} are i.i.d. $N(0, \sigma^2)$ measurement errors, $i = 1, \cdots, n_k$. For comparison of treatment effects, a hypothesis of homogeneity H_0: $\mu_0 = \mu_1 = \cdots = \mu_K$ versus H_1: at least two μ_k's are not the same, can be tested using an ANOVA F-test.[23] More specific tests using linear contrasts can also be considered.[24] Note that with the ANOVA approach no assumption is made on the shape of the DR.

A more elegant approach is to consider a regression model for Y_{ik}'s, with dose level as a continuous predictor:

$$Y_{ik} = \mu(d_k, \theta) + \varepsilon_{ik} (i = 1, \ldots, n_k; k = 0, 1, \ldots, K), \quad (6)$$

where $\boldsymbol{\theta}$ is a vector of model parameters, $\mu\left(d_k, \boldsymbol{\theta}\right)$ is the mean response at dose level d_k, and ε_{ik} are i.i.d. $N(0, \sigma^2)$ measurement errors. There are many commonly used functional forms for $\mu\left(d, \boldsymbol{\theta}\right)$ and one popular choice is a four-parameter sigmoid E_{max} model.[20] A nonlinear least squares estimate of $\boldsymbol{\theta}$ can be obtained by minimizing the deviance $\sum_{k=0}^{K} \sum_{i=1}^{n_k} \left(Y_{ik} - \mu\left(d_k, \boldsymbol{\theta}\right)\right)^2$. A significance test of the DR relationship can be carried out by testing the significance of a certain coefficient in the model.[25] An advantage of DR model in Equation (6) over the ANOVA model is that by estimating the parameters $\boldsymbol{\theta}$, one can obtain estimates of response mean and variance at doses that were not evaluated in the experiment. A potential disadvantage is the difficulty of modeling bias due to model misspecification.

One way of handling model uncertainty is a combination of multiple comparisons with modeling techniques (MCP-Mod).[26] With MCP-Mod, a set of plausible DR models is prespecified at the trial design stage. When experimental data become available, the significance of the DR is tested for each individual model with proper adjustment for multiplicity. If the overall null hypothesis of a "flat" DR cannot be rejected, the procedure stops; otherwise, statistically significant models are considered to estimate the parameters of interest. MCP-Mod received a positive qualification opinion from the Committee for Medicinal Products for Human Use in 2014[27] and a fit for purpose determination from the US Food and Drug Administration (FDA) in 2016.[28] It provides a very good success story of collaboration among statisticians, PMx scientists, and other researchers across industry, academia, and regulatory bodies.[29]

The DR model in Equation (6) assumes the same average response for each individual treated at a given dose. In reality, DR can vary across individuals due to intrinsic variability in their drug exposure. The ER model may be a more informative approach to the problem. For illustration supposes, the exposure is represented by the steady-state area under the concentration curve (AUCss), and the response is a single measurement at some prespecified time point at steady-state. Let CL_{ik} denote the apparent clearance of the ith subject in the kth dose group, in which case $\mathrm{AUCss}_{ik} = d_k/\mathrm{CL}_{ik}$. Acknowledging variability of clearance in the patient population, the individual clearance can be modeled as $\log \mathrm{CL}_{ik} \sim N(\log \mathrm{TVCL}, \omega^2_{\mathrm{CL}})$, where TVCL denotes the typical value of clearance (e.g., population median value of clearance), and ω_{CL} is the between-subject SD of clearance. Suppose a four-parameter $E_{\max}$ model is a plausible description of the ER relationship. Then, conditional on AUCss_{ik} (derived via CL_{ik} estimates coming from the PK model), an ER model can be written as:

$$Y_{ik} = E_0 + \frac{E_{\max} \mathrm{AUCss}^h_{ik}}{EC^h_{50} + \mathrm{AUCss}^h_{ik}} + \varepsilon_{ik}, \quad (7)$$

where $(E_0, E_{\max}, \mathrm{EC}_{50}, h)$ are model parameters and ε_{ik} are i.i.d. $N(0, \sigma^2)$ measurement errors. The ER model (in Equation (7)) takes into account inter-individual variability in drug exposure, thereby providing a more thorough description of the drug effect relationship than the DR model (in Equation (6)). A limitation of the ER model is that if only sparse PK sampling is done on all participants, then there could be issues of model misspecification. However, this problem can be mitigated by taking intensive PK samples

from a subset of participants for model identification or model confirmation, and then taking sparse PK samples on the rest.

More examples of ER models, as well as different modeling and simulation strategies can be found in recent books.[30–32]

INTEGRATION OF FIELDS

Although statistics has traditionally been regarded as a core discipline in drug development and regulatory decision making, the value of PMx has been increasingly recognized, especially over the past 2 decades.[6,7,9] In 2016, the European Federation of Pharmaceutical Industries and Associations (EFPIA) introduced the term "Model Informed Drug Discovery and Development" (MID3) signifying that R&D decisions are "informed" rather than "based" on model-derived outputs.[33,34] MIDD can be viewed as a subset of MID3 that deals with the clinical development stage. Recently, the FDA launched the MIDD pilot program to engage biopharmaceutical developers in early interaction with the FDA on clever use and application of MIDD to optimize their clinical development programs.[35]

The key ingredients of MIDD are described elsewhere.[5,7,9] In what follows, we discuss some important applications of MIDD that create special opportunities for collaboration and synergy of statisticians, pharmacometricians, and other quantitative scientists in drug development. Fundamentally, the MIDD strategy should start with an understanding and formulation of the key research questions and the quantification of the target values and decision making criteria.[9] Subsequent choice of analytic tools in the pursuit of the optimal solution is the joint task of a statistician and a pharmacometrician.

PK/PD modeling and simulation

The PK/PD modeling process was briefly described in the "Pharmacometric Approach" section. The PD component may represent a measure of safety and/or efficacy, and it can be a continuous, categorical or binary outcome, depending on the research context. PK/PD modeling and simulation (M&S) provide a powerful toolkit that is applied continuously throughout the R&D process. For instance, based on preclinical (animal) studies, PK/PD models are built to calibrate the dose range for phase I first-in-human studies. Thereafter, phase I safety, PK, and PD data are utilized to develop PK/PD models to project the doses and regimen to achieve the desired effect on relevant biomarkers and early clinical efficacy in phase IIa proof-of-concept studies. Furthermore, PK/PD models are updated to optimize dosing and sampling schemes in phase IIb dose-ranging studies in target patient populations and, later, assess the probability of success of phase III pivotal studies and assess the need for dose adjustments in special populations. A paper by Kowalski 7 provides some excellent real-life examples of using PK/PD M&S to develop quantitative go/no-go decision criteria and the pivotal role of statisticians and PMx scientists in this process.

Disease progression models

Disease progression models use mathematical relations to quantify the time course of the disease with respect to some relevant biomarkers or clinical end points.[36] Integration of disease models with PK/PD models can further help evaluate the impact of drug on the disease trajectory; for instance, to assess how a particular drug intervention can slow down the disease progression. One can distinguish three types of disease progression models: empirical; semimechanistic; and systems biology.[36] Empirical models are frequently applied in diseases where underlying mechanisms are elusive (e.g., neuropsychiatry) and the clinical outcomes are subjective scores (e.g., patient-reported outcomes or clinician's assessments). By contrast, systems biology models provide comprehensive descriptions of the biological and/or molecular pathways (e.g., bone remodeling in osteoporosis). Semimechanistic models may utilize knowledge of both the biology and pathophysiology of the disease, but in a less comprehensive way than systems biology models. It stands to reason that both statistical and PMx expertise is essential for implementing such models.

There are several merits of disease progression models in drug development. First, they can allow prediction of individual disease trajectories, taking into account relevant covariates and uncertainties, thereby helping investigators to identify eligible subjects for clinical trials. In addition, these models can be used to optimize clinical trial designs through M&S, and they can frequently yield more powerful statistical analyses than standard statistical approaches, such as linear MMRMs.[37] Disease models are now widely adopted in drug development. One recent example is the regulatory endorsement of the clinical trials simulation tool for Alzheimer's disease progression, both by the FDA and the European Medicines Agency (EMA).[38]

Model-based meta-analysis

Model-based meta-analysis (MBMA) provides a multidimensional framework to integrate available knowledge

to quantitatively address important research questions, facilitate drug development decisions, and promote lifecycle management activities.[39] With MBMA, we build some longitudinal or multivariate meta-regression models to characterize drug effects (on a class or individual agent level), dose and regimen effects, patient population effects, etc. Both individual patient-level data and trial summary-level data can be aggregated, thereby embracing the "totality of evidence" paradigm to address the questions of interest. Unlike traditional meta-analyses that either combine simple summary results or individual participant data from several RCTs, MBMA applies more advanced scientific modeling, based on drug characteristics and the biology of the disease, while accounting for heterogeneity of RCTs.

MBMA can be utilized across the entire R&D continuum.[39] In particular, MBMA can help to rigorously establish safety and efficacy targets needed for compound differentiation; to design cost-effective clinical trials; to iteratively quantify probability of success of a clinical development program; to leverage information for drugs with shared mechanisms of action, etc. Examples of successful applications of MBMA are numerous across different indications, such as rheumatoid arthritis,[40,41] atrial fibrillation,[42] multiple myeloma,[43] and idiopathic pulmonary fibrosis,[44] just to name a few. MBMA requires synthesis of knowledge from various data sources and calls for close collaboration of cross-functional drug development teams.

Bayesian and PK/PD-driven trial designs

Phase I first-in-human studies are conducted to explore safety, tolerability, and PKs of the compound, and to identify the maximum tolerated dose (MTD). Most such studies are cast as adaptive dose-escalation designs where subjects are assigned sequentially or in cohorts to increasing dose levels, provided that previous doses are deemed as sufficiently safe.[45] Statisticians have played the central role for calibrating phase I designs. Incorporating PMx (PK/PD models) in dose-escalation decisions is not very common; yet, some authors demonstrated that this approach may improve design efficiency.[46–48] One recent paper[49] described a systematic comparison of several phase I Bayesian adaptive designs utilizing PK data in dose escalation. It was found that, for the studied scenarios, trial efficiency (as measured by the number of observed dose-limiting toxicities or the probability of correct MTD selection) is not improved, but the estimation of the dose–toxicity curve may be better compared to the designs that do not utilize PK data. Another recent paper[50] proposed a Bayesian dose finding design using dose-exposure data in the escalation rule and found that it may improve the probability of correct dose selection. An important and potentially useful consideration is the assessment of distributed designs (evaluating a range of doses) versus concentrated design (testing only placebo and the maximum dose). Comparing these strategies via simulation may help investigators select the "best" design option for the chosen objectives and decision making criteria, as was demonstrated in the context of a first-in-patient study in psoriasis.[51]

Another great opportunity for fusion of statistical and PMx ideas is in the design of phase II dose-ranging studies.[25,52] For instance, MCP-Mod[26] is very useful and widely applied in drug development. More recently, PMx extensions of MCP-Mod have been proposed.[53,54]

Optimal designs

Optimal designs (ODs) present an example of a methodology where statistical and PMx approaches can provide further interdisciplinary advances.[55] For a PK/PD experiment with a given NLMEM, one can construct a likelihood function and obtain the Fisher Information Matrix (FIM), which depends on the model parameters and the design points. ODs are obtained by maximizing some optimality criteria of the FIM (e.g., the D-optimal design maximizes the determinant of the FIM). ODs for exposure-response models can potentially result in significant improvement in the accuracy of estimates compared to more standard designs.[56] In most practical applications, ODs are so-called locally optimal (i.e., they depend on model parameters that are unknown upfront). To implement them in practice, one can perform pilot studies or elicit from experts' preliminary estimates or nominal values for the unknown parameters. Bayesian ODs and minimax (maximin) ODs are important generalizations of locally optimal designs.

Bayesian ODs are useful when experts or pilot studies provide different nominal values of the parameters, and the experimenter has varying level of confidence in these values. Minimax or maximin ODs focus on optimization of the chosen criterion under the most misspecified nominal values of the parameters.[57,58] In sequential experiments where responses are available at a short time, one can consider implementing model-based adaptive optimal designs (MBAODs).[59,60] Although MBAODs can be highly efficient, they are very challenging computationally. Due to the complex structure of the design space, common gradient-based methods are generally not helpful, and more advanced optimization techniques, such as global search algorithms, are required. Nature-inspired metaheuristic algorithms are general purpose optimization tools that do not require assumptions for them to work

well, and they may provide high-speed efficient computational methodologies.[61–63]

PK bridging approaches

Use of PK/PD modeling is of paramount importance for therapeutic development in special populations, such as pediatric, geriatric, patients with rare diseases, etc. For instance, in pediatric studies, one typically has prior information on PK and PD of the drug in adults. However, exposing children to the same dosing regimen as adults may lead to severe overexposure and serious adverse effects. Therefore, having a well-defined dose—exposure—response relationship in the pediatric target population is essential to allow for a treatment which is both safe and efficacious. Extrapolating relevant information from adult studies through PK/PD modeling and through Bayesian approaches in this context provide a great opportunity for collaboration between statisticians and pharmacometricians.[64,65] Health authorities are fully supportive of such innovative approaches.[66,67] Some good examples of applications of PMx and statistical methodologies in pediatric clinical development have been documented.[68–70] Many other challenging problems, such as bridging development from preclinical (animal) experiments to studies in humans, bridging different drug formulations or routes of administration, estimation of equipotent doses of drugs with similar mechanisms of action[71] can be tackled using similar approaches.

Biomarker-based development and precision medicine

The precision medicine paradigm is focused on tailoring treatment decisions to an individual (i.e., finding the right treatment for the right patient at the right time).[13] The Precision Medicine Initiative announced by President Barak Obama in 2015 has received full recognition and endorsement over the past few years. In 2019, the FDA issued guidance on enrichment strategies for clinical trials of drugs and biologics,[72] affirming the importance of new precision medicine product development. New statistical methodologies, such as dynamic treatment regimes,[73] adaptive enrichment designs,[74] and master protocols[75] have been emerging and receiving fast uptake in practice. PMx methods also hold great promise in precision medicine, as the goal of personalizing treatment is in good alignment with the idea of population PK/PD modeling. One recent paper[76] describes emerging roles of clinical PMx in cancer precision medicine and provides a roadmap for future precision oncology with challenges and opportunities for interdisciplinary collaboration. Another useful idea for precision medicine is a dashboard system that integrates real-time patient monitoring and calculation of optimal dose and timing of dose delivery.[77] An example of a dashboard system is a fully automated artificial pancreas system for patients with type 1 diabetes.[78] Development and validation of dashboards for precision dosing requires interdisciplinary efforts combining PMx, optimization, machine learning, and big data analytics.

A perspective for the future

A recent paper 10 provides a strategic perspective on the evolution of PMx and quantitative systems pharmacology for the next decade. The authors anticipate these fields to continue profoundly impact drug development and use. We concur with this vision and additionally assert that statistics, data science, and machine learning will also be essential ingredients to help increase the efficiency of the R&D enterprise.

A SYNERGY BETWEEN PHARMACOMETRICS AND STATISTICS

In this section, we provide three examples to illustrate the value of incorporating ideas from statistical and PMx approaches to tackle complex problems in clinical drug development. The importance of collaboration between statistical and PMx scientists to expedite drug discovery and development cannot be overemphasized.

Exposure-response modeling and randomization-based inference

It is increasingly recognized that DR and ER analyses can increase power and improve cost-efficiency of clinical trials.[79,80] However, these analyses invoke population model assumptions (such as normal random sampling), and if these assumptions are violated, statistical estimators and tests may be biased. On the other hand, randomization-based tests ensure valid inference (i.e., a prespecified probability of a false positive finding can be achieved but not exceeded).[18] However, randomization-based tests are typically very simplistic in that they ignore important information on drug exposure. At a glance, ER modeling and randomization-based inference are disjoint concepts. Here, we illustrate how their combination can improve study power while maintaining validity of the test.

For illustration purposes, we consider a hypothetical phase II proof-of-concept clinical trial, designed as a randomized, parallel group 3-arm study to evaluate the effects of a low dose (50 mg), high dose (100 mg), and placebo (0 mg) with respect to a continuous outcome. The study objective is to obtain evidence of the drug effect compared to placebo.

We apply a recently published roadmap to randomization in clinical trials[81] to evaluate different combinations of design and analysis strategies via simulation. We investigate three important aspects in this setting: (1) choice of a randomization procedure; (2) choice of a data analytic approach (ANOVA model vs. ER model); and (3) decision on whether to ignore or include randomization in the analysis (conventional likelihood-based approach vs. randomization-based inference). The criteria for assessing the merits of a particular strategy are validity (type I error rate) and efficiency (power).

The full simulation details on the data generating mechanism and the data analysis strategies are presented in the Supplementary Appendix. In essence, we consider two different randomization procedures—random allocation rule (Rand) and truncated multinomial design (TMD)—to randomize a total of $n = 24$ subjects equally among three treatment groups (8 per arm).

We assume that the drug exposure at the site of action is causal in driving the drug response, and the true ER relationship is linear that is, the mean response conditional on the ith subject's exposure at steady state is given by:

$$f\left(\text{AUCss}_i, \beta_0, \beta_1\right) = \beta_0 + \beta_1 \text{AUCss}_i. \quad (8)$$

With these assumptions, the response of the ith patient, Y_i, is generated as follows:

$$Y_i = f\left(\text{AUCss}_i, \beta_0, \beta_1\right) + u_i + \varepsilon_i, \quad i = 1, \ldots, 24. \quad (9)$$

In Equation (9), there are three components that account for the response of the ith patient:

- Mean response at the steady-state exposure level; see Equation (8). We investigate two scenarios: (A) Flat ER: $\left(\beta_0, \beta_1\right) = (5, 0)$, and (B) Non-flat linear ER: $\left(\beta_0, \beta_1\right) = (5, 0.7)$.
- u_i is an unknown term associated with the patient (e.g., this can be an important unknown prognostic covariate affected by a time trend). We investigate two cases: (I) No time trend: $u_i \equiv 0$ for $i = 1, \cdots, 24$, and (II) Linear trend: $u_i = i/5$ for $i = 1, \cdots, 24$, which indicates that patients enrolled later in the trial tend to have higher values of the response.
- ε_i are measurement error terms, assumed to be i.i.d. random variables. We explore two choices for the distribution of ε_i: (i) Standard normal: $N(0, 1)$, and (ii) Cauchy: $C(0, 1)$.

In all, we have four dimensions to the data generating mechanism: (1) randomization procedure (Rand or TMD); (2) ER relationship (flat or non-flat); (3) linear trend (yes or no); and (4) measurement error distribution (standard normal or Cauchy).

For the data analysis, we consider four methods that are determined by the modeling strategy (ANOVA or ER model), and whether or not randomization design is accounted for in the analysis (No ⇒ likelihood-based inference, or Yes ⇒ randomization-based inference). More specifically:

1. ANOVA model, likelihood-based inference. The responses Y_{ik} ($i = 1, \cdots, 8$ and $k = 1, 2, 3$) are assumed to be independent normal variables with $E\left(Y_{ik}\right) = \mu_k$ and $\text{var}\left(Y_{ik}\right) = \sigma^2$. A null hypothesis $H_0: \mu_1 = \mu_2 = \mu_3$ is tested against the alternative H_1: at least two treatment means are not the same using ANOVA F-test.[23]
2. ER model, likelihood-based inference. Given observed exposure and response data, we fit a linear model $E\left(Y_i\right) = \beta_0 + \beta_1 AUCss_i$ ($i = 1, \cdots, 24$), obtain the least squares estimates $\left(\widehat{\beta}_0, \widehat{\beta}_1\right)$, and test significance of the ER relationship by testing the null hypothesis of a zero slope ($H_0: \beta_1 = 0$).
3. Nonparametric ANOVA, randomization-based inference. Under the null hypothesis of equality of treatment effects, the outcome of each subject is not affected by assigned treatment. To quantify evidence against the null, treatment assignments are permuted in all possible ways consistent with the randomization design, and the randomization p value is the sum of null probabilities of the treatment assignment permutations in the reference set that yield the test statistic values greater than or equal to the experimental value. We apply a nonparametric rank-based Kruskal-Wallis test[82] as a measure of between-group difference and estimate a randomization-based p value via Monte-Carlo simulation.[83]
4. ER model, randomization-based inference. If ER is flat, individual outcomes are not affected by drug exposures. We permute treatment assignments (and corresponding values of exposures; i.e., the AUCss_i values) in all possible ways consistent with the randomization design. For each permutation, we fit a linear model $E\left(Y_i\right) = \beta_0 + \beta_1 \text{AUCss}_i$, $i = 1, \cdots, 24$ and perform the test of significance of the slope ($H_0: \beta_1 = 0$) by calculating the t-statistic $t = \widehat{\beta}_1 / SE(\widehat{\beta}_1)$. A randomization-based p value is the sum of null probabilities of the treatment assignment permutations in the reference set that yield

randomization sequences with a t-statistic at least as extreme as the experimental value. The Monte-Carlo simulation is used to estimate the randomization-based p value.

To our knowledge, the latter approach (ER modeling followed by a randomization-based analysis) has not been explored previously. Permutation tests and resampling procedures have been considered in the PMx modeling context[84–86]; however, they were not reflective of a specific randomization procedure. By contrast, our described approach directly uses the reference set of a chosen randomization procedure as a basis for statistical test of significance using an ER model.

The simulations were performed using R (https://www.r-project.org/) and Julia (https://julialang.org) programming languages. For each combination of experimental scenario and data analysis strategy, a clinical trial with $n = 24$ patients was simulated 10,000 times. All statistical tests were two-sided, using significance level of $\alpha = 0.05$. The type I error rate (when ER was flat) and power (when ER was non-flat) were calculated as the proportion of simulation runs for which the chosen test exhibited a statistically significant result ($p < 0.05$).

Figure 1 summarizes the key findings from our simulations. There are two columns (flat ER and non-flat ER), which show type I error rate and power, respectively. In addition, there are two rows corresponding to different randomization designs (Rand and TMD). Within each of the four plots, there are three scenarios: "normal" (i.e., when the normal random sampling assumption is met); "time trend" (i.e., when there is a time drift in the study population); and "Cauchy" (i.e., when the measurement error distribution is misspecified and heavy-tailed). For each scenario, we have four data analytic strategies, as described above.

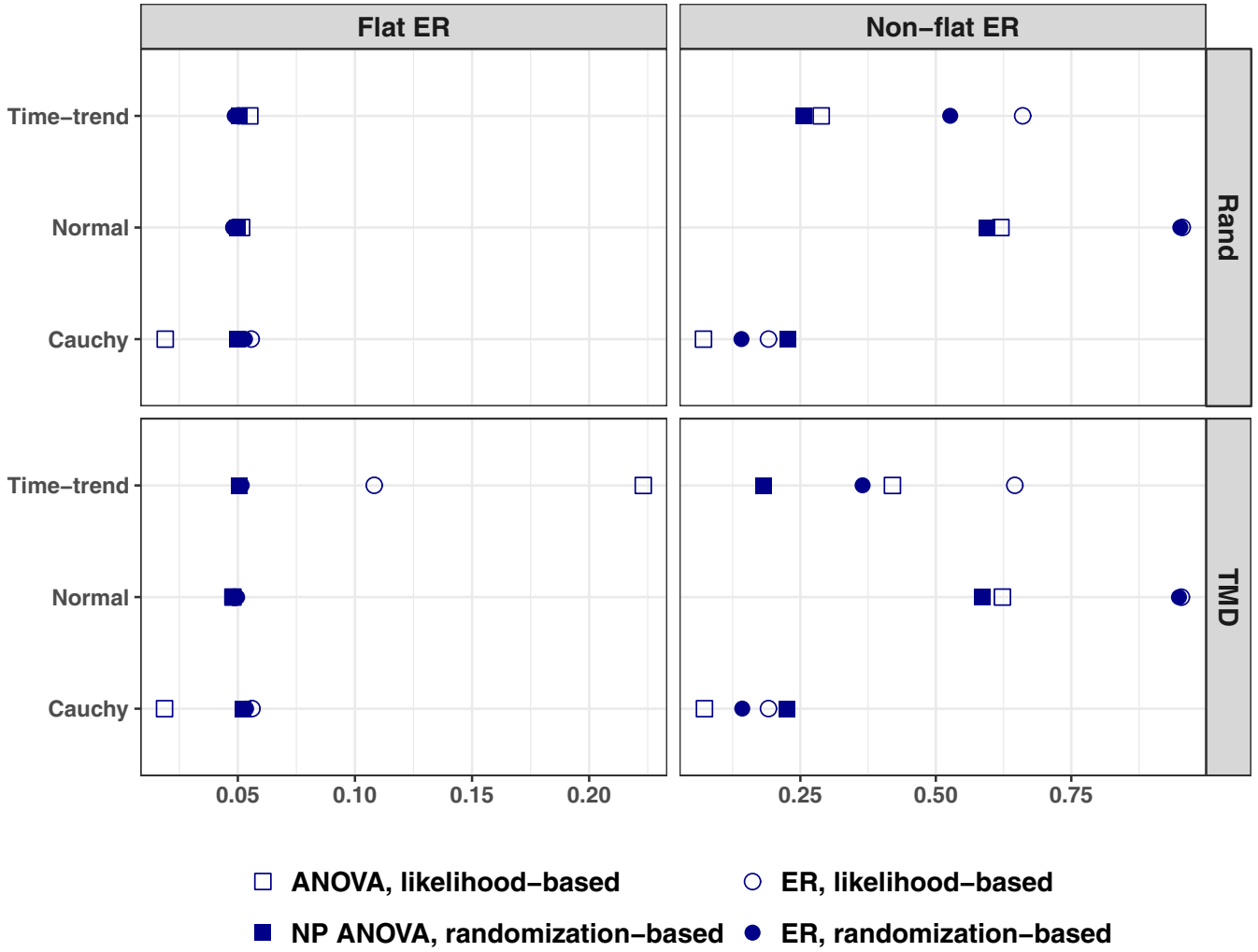

FIGURE 1 Simulated type I error rate (flat ER) and power (non-flat ER) for two randomization designs (Rand and TMD), under three experimental scenarios (time trend, normal, and Cauchy), and four data analytic strategies. ANOVA, analysis of variance; ER, exposure–response; Rand, random allocation rule; TMD, truncated multinomial design.

Under the "normal" scenario, the type I error rate is maintained at the 5% nominal level for either Rand or TMD, for all four data analytic approaches. In regard to power, the two ER approaches have substantially higher power (~95%) than the likelihood-based ANOVA (~62%) and the randomization-based NP ANOVA (~59%).

Under the "time trend" scenario, the designs exhibit more differential performance. First, consider the "flat ER" case. For the Rand design, all four analytic methods maintain the type I error rate at 5%. By contrast, for the TMD design, the randomization-based methods do maintain the type I error rate at 5%, but the likelihood-based methods have inflated type I error: ~ 22% for the ANOVA model, and ~ 11% for the ER model. Inflation of type I errors of some conventional tests following TMD under linear time trends has been noted previously.[87,88] Next, let us consider the "non-flat ER" case. For the Rand design, the power of the two ER approaches are ~ 66% for the likelihood-based ER, and ~ 53% for the randomization-based ER, whereas the power of the two ANOVA approaches is only 26–29%. For the TMD design, the most powerful approach is likelihood-based ER model (~ 65% power) and the randomization-based NP ANOVA model has the lowest power (~ 18%). Overall, under time trend, statistical power is reduced compared to the normal random sampling scenario, and Rand is more powerful than TMD.

For the "Cauchy errors" scenario, the results are consistent for both the Rand and TMD designs. Under the "flat ER," the type I error rate is maintained at 5% for the two randomization-based approaches, and it is ~ 2% for the likelihood-based ANOVA and ~ 6% for the likelihood-based ER model. Under the "non-flat ER," the most powerful approach is randomization-based NP ANOVA (~ 22% power) and the least powerful one is likelihood-based ANOVA (~ 7% power). This speaks to the importance of having a robust analytic procedure in the presence of model misspecification and outliers.[53]

Secukinumab clinical development success story

In the literature, there are many examples of successful applications of MIDD, showing added value from various stakeholders' perspectives, such as cost savings, improved quality of decision making, reduced trial burden for vulnerable populations, accelerated access for patients to innovative therapies, simplified treatment posology, etc.[89,90] Here, we briefly describe one clinical development success story.

Secukinumab (COSENTYX) was developed as a first-in-class fully human monoclonal antibody against the proinflammatory cytokine interleukin-17A (IL-17A) that plays a key role in the pathogenesis of plaque psoriasis. The secukinumab clinical program for moderate to severe psoriasis (that formed the basis for registration in 2014) included 10 phase II and phase III studies. The phase II program consisted of a proof-of-concept study with a single i.v. dose of 3 mg/kg, followed by three dose-ranging and regimen-finding studies whose designs were informed by modeling and combinatorial optimization of route of administration, dose, and regimen. Importantly, in both phase II and phase III studies, standard efficacy end points, such as Psoriasis Area and Severity Index 75 response (PASI 75) and Investigator's Global Assessment (IGA) for clear to almost clear skin, were used. A hallmark of the program was comprehensive M&S using data from phase II studies to calibrate two dose regimens (150 mg or 300 mg administered subcutaneously at weeks 0, 1, 2, and 3 followed by monthly maintenance dosing) as most promising from the risk-benefit perspective for testing in pivotal phase III studies. Four double-blind, randomized, placebo-controlled phase III studies confirmed clinical efficacy and acceptable safety profiles of both regimens, which had not been tested in phase II. The 300 mg regimen was recommended as the optimal clinical dose.

This success story highlights several important takeaway messages. First, it shows the value of the MIDD approach, which includes iterative model building and reviewing theory and assumptions in light of accumulating clinical trial evidence. Second, it taps on the power of continuous collaborative efforts of statistical and PMx scientists. From the second author's personal communication with both statistical and PMx leads of the secukinumab program, some key ingredients of collaboration included: working with your counterpart in mind, working together (and not in parallel), understanding, discussing, and constructively challenging self and others. Third, there is a sense of responsibility and joint ownership of the deliverables, irrespective of the line functions within the company. Quoting Harry S. Truman (the 33rd president of the United States), "it is amazing what you can accomplish if you do not care who gets the credit."

Model-based adaptive optimal design for pediatric PK bridging studies

As mentioned in "PK Bridging Approaches" section, in pediatric studies, one typically has prior information about the PK/PD of the drug in adults. Even if it is expected that, given the same exposure in children compared with adults, the effect of the drug will be the same, studies to demonstrate that similar exposures can be

reached are typically required (so-called PK matching studies). However, achieving these similar PK exposures can be challenging given that the PK may differ greatly between children and adults due to physiological differences, such as body size and renal maturation. Further, determining the number and ages of the children to include in the study may be challenging.

Wang et al.[91] proposed a precision criterion for use in sample size determination when designing pediatric PK studies. This precision criterion requires 95% confidence intervals around CL and V to be within 60% to 140% of the geometric mean value of the parameters in all included age groups of children with at least 80% power. To meet this criterion, CL and V can be computed using nonparametric analyses that are reliant on rich data in a relatively high number of patients or model-based approaches, where sparser data and smaller studies may be possible, with the potential risk of model misspecification. For planning purposes, one can take estimates of the SE of CL and V from a variety of different sources.[92] However, a misspecified guess of initial parameters and variability may lead to an inaccurate sample size and suboptimal design. Statistically grounded MBAODs are expected to be less sensitive than nonadaptive methods to misspecification in the design stage.[59,93,94]

A standard pediatric scaling model, as suggested by Germovsek et al.,[95] to describe differences in PKs between adults and children may assume that CL and V are scaled according to size (allometric scaling, based on weight [WT]) and that CLs are also scaled according to maturation (based on post menstrual age [PMA]). For example, the PK model for CL_i and V_i might have the following form:

$$\mathrm{CL}_i = \beta_{\mathrm{CL}} \cdot e^{\eta_{\mathrm{CL},i}} \cdot \left(\frac{\mathrm{WT}_i}{70}\right)^{0.75} \cdot \frac{\mathrm{PMA}_i^{\beta_{\mathrm{Hill}}}}{\mathrm{PMA}_i^{\beta_{\mathrm{Hill}}} + \beta_{\mathrm{TM50}}^{\beta_{\mathrm{Hill}}}}$$

$$V_i = \beta_V \cdot e^{\eta_{V,i}} \cdot \left(\frac{\mathrm{WT}_i}{70}\right)$$

Typically, the maturation function of the CL_i function will be at a maximum (full maturation) sometime in infancy. If this model describes the reality, one can clearly see that variability will be larger for infants than for adults, indicating that more subjects at lower age ranges will be required to meet the criterion proposed by Wang et al.[91] This also indicates that the standard statistical method of designing pediatric PK studies will be underpowered in infant age groups, where the variability seen in adult data is typically assumed to be the same in children/infant PK data.

In a simulation study,[59] concentrations of two different theoretical drugs for an adult study population were simulated using the above model. Once the values of variability of CL_i and V_i were estimated, they were used in different ways to determine sample sizes of pediatric trials using both standard statistical methods and PMx approaches. For the latter, the maturation function was not identifiable via adult data, so parameter values had to be assumed for a priori design planning and those guesses were made to be either close to the simulated truth or misspecified. Simulations of these planned pediatric trials were then performed and the percent of those trials that, in the end, achieved the precision criteria,[91] were computed. Included in the comparison was a number of MBAOD approaches, that allowed for updates to the distribution of sample sizes within different age groups of children after interim analyses, an early stopping criterion if the precision criteria were achieved at an interim analysis, and adjustments to the process to account for multiple comparisons. Results showed that MBAODs (combining both statistical and PMx ideas) required fewer children, on average, to fulfill a precision criterion than the sample size obtained from the more traditional estimation methodologies.[59]

DISCUSSION

Drug development has been undergoing a major paradigm shift—from a sequence of distinct phases toward a more continuous, MIDD paradigm that integrates relevant knowledge to optimize clinical development decisions. In this paper, we provided examples of applications of statistical and PMx approaches, as well as their combinations, which can yield synergy for solving complex problems in clinical drug development. Table 1 summarizes the three examples presented in the section "A Synergy Between Pharmacometrics and Statistics".

With evidence generation now increasingly being based on data accrued both within and outside of a clinical development program, integrative approaches in translational medicine that are iterative in nature and demand commitment to a "totality of evidence" mindset will be crucial for success.[96–98] To realize the promise of MIDD fully, statistics and PMx will need to operate more in partnership than as complementary fields that generate their own sets of analyses. They will need to collectively own the problem space, define the right questions, and arrive at the appropriate designs of trials or analysis plans to help answer the pertinent questions. Both statisticians and PMx scientists possess unique qualifications that can contribute better to the research enterprise. Statisticians have the depth and breadth of knowledge of statistical principles and frameworks and how these can address specific research questions. PMx scientists have the depth of knowledge of the dynamics between biological models and the

TABLE 1 Added values of statistical and pharmacometric approaches in the considered examples in "A Synergy Between Pharmacometrics and Statistics" section

Example	Added value of the statistical approach	Added value of the pharmacometric approach	Synergy/implication
"Exposure-Response Modeling and Randomization-Based Inference" section	Randomization in the design and analysis	Exposure-response modeling	Valid, more robust and more powerful test
"Secukinumab Clinical Development Success Story" section	Choice and implementation of proper design and analysis of phase II and phase III RCTs	Knowledge integration; M&S to predict optimal dose regimens for phase III	Validation of model-informed dose regimens in phase III RCTs
"Model-Based Adaptive Optimal Design for Pediatric PK Bridging Studies" section	Proper handling of multiple comparisons; handling of adaptive designs; optimal design techniques	Population PK modeling with maturation and size scaling; potential sample size reductions	Sample size reductions with appropriate handling of statistical tests. Optimization and adaptation to improve estimation and reduce modeling bias.

Abbreviations: M&S, modeling and simulation; PK, pharmacokinetic; RCT, randomized controlled trial.

pharmacology of medical agents under investigation. Moreover, with the increasing pace of innovations in each of their respective fields, statisticians and PMx scientists should continue to correspond on methodological questions, as this can also result in "hybrid" innovations.[26,54]

In this paper, we have made an attempt to demonstrate the value of integration of ideas from statistics and PMx, the need for applying rigorous scientific judgment when evaluating different approaches, and the need for collaborative work in drug development. Different approaches to solving complex problems exist, and novel ones, such as data science and machine learning, are emerging rapidly, which may call for a three-way interaction of statisticians, PMx scientists, and data scientists.[99] However, it is not the approach itself that we should primarily focus on, but rather the important drug development problems we are trying to solve. Here, we are in good concordance with Lewis Sheiner's position[100]:

> "...good statistics are absolutely essential to good clinical investigation, and hypothesis testing, when used judiciously and appropriately, can be a useful inferential tool. What is wrong is that a particular statistical practice (and its associated epistemologic view) has become almost mandatory, to be applied willy-nilly to drug trials, regardless of the purposes they are meant to serve. All thoughtful scientists, including statistical scientists, should join me in rejecting this. The intellectual illness of clinical drug evaluation that I have discussed here can be cured, and it will be cured when we restore intellectual primacy to the questions we ask, not the methods by which we answer them."

In summary, statistics and PMx definitely have more in common than what keeps them apart. Researchers from both disciplines should find ways to collaborate and synergize to achieve the common and important goal of improving clinical research and development, and delivering novel and efficacious medicines to patients with medical need.

ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers for their constructive critical remarks, which helped significantly improve the content and the presentation of the paper. The second author is grateful to his Novartis colleagues – Frank Bretz, Brian P. Smith, Oliver Sander, Achim Guettner, and Thomas Dumortier for their feedback on the manuscript and their help to build the secukinumab clinical development success story example in "Secukinumab Clinical Development Success Story" section. G.M.'s work was supported by the Statistical and Data Management Center of the International Maternal Pediatric Adolescent AIDS Clinical Trials Network under the National Institute of Allergy and Infectious Diseases grant No. UM1 AI068616.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

Conception: OS and YR. *Writing of the manuscript*: OS, with contributions from YR, EMS, GM, ACH, and WKW. *Design of simulation studies*: OS and YR. *Development of code and running simulations*: YR. All authors reviewed and edited both the original manuscript and the revised version. The authors read and approved the final version.

ORCID

Yevgen Ryeznik https://orcid.org/0000-0003-2997-8566
Oleksandr Sverdlov https://orcid.org/0000-0002-1626-2588
Elin M. Svensson https://orcid.org/0000-0002-0093-6445
Grace Montepiedra https://orcid.org/0000-0002-6866-4335
Andrew C. Hooker https://orcid.org/0000-0002-2676-5912
Weng Kee Wong https://orcid.org/0000-0001-5568-3054

REFERENCES

1. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Economics*. 2016;47:20-33.
2. ICH E9: Statistical principles for clinical trials. 1998. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf
3. Williams PJ, Ette EI. Pharmacometrics: impacting drug development and pharmacotherapy. In: Ette, EI. & Williams, PJ., eds. *Pharmacometrics: The Science of Quantitative Pharmacology*. John Wiley & Sons, Inc.; 2007:1-21.
4. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther*. 1997;61:275-291.
5. Lalonde RL, Kowalski KG, Hutmacher MM, et al. Model-based drug development. *Clin Pharmacol Ther*. 2007;82:21-32.
6. Milligan PA, Brown MJ, Marchant B, et al. Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development. *Clin Pharmacol Ther*. 2013;93:502-514.
7. Kowalski KG. Integration of pharmacometric and statistical analyses using clinical trial simulations to enhance quantitative decision making in clinical drug development. *Stat Biopharm Res*. 2019;11:85-103.
8. Senn S. Statisticians and pharmacokineticists: what they can still learn from each other. *Clin Pharmacol Ther*. 2010;88:328-334.
9. Kowalski KG. My career as a pharmacometrician and commentary on the overlap between statistics and pharmacometrics in drug development. *Stat Biopharm Res*. 2015;7:148-159.
10. Mentré F, Friberg LE, Duffull S, et al. Pharmacometrics and systems pharmacology 2030. *Clin Pharmacol Ther*. 2020;107:76-78.
11. Chatfield C. Model uncertainty, data mining and statistical inference. *J Royal Stat Soc Series A*. 1995;158:419-466.
12. Lee JJ, Chu CT. Bayesian clinical trials in action. *Stat Med*. 2012;31:2955-2972.
13. FDA Precision Medicine. (2018). https://www.fda.gov/medical-devices/in-vitro-diagnostics/precision-medicine
14. Mallinckrodt CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Inform J*. 2008;42:303-319.
15. Armitage P. The role of randomization in clinical trials. *Stat Med*. 1982;1:345-352.
16. Rosenberger WF, Lachin JM. *Randomization in Clinical Trials*. John Wiley & Sons Inc; 2016.
17. Fisher RA. *The Design of Experiments*. Oliver and Boyd; 1935.
18. Rosenberger WF, Uschner D, Wang Y. Randomization: the forgotten component of the randomized clinical trial. *Stat Med*. 2019;38:1-12.
19. Gibaldi M, Perrier D. *Pharmacokinetics*. Informa Healthcare; 1982.
20. Macdougall J. Analysis of dose-response studies -- Emax model. In: Ting, N, eds. *Dose Finding in Drug Development*. Springer; 2006:127-145.
21. Bauer RJ. NONMEM tutorial part I: description of commands and options, with simple examples of population analysis. *CPT Pharmacomet Sys Pharmacol*. 2019;8:525-537.
22. Bauer RJ. NONMEM tutorial Part II: estimation methods and advanced examples. *CPT Pharmacomet Syst Pharmacol*. 2019;8:538-556.
23. Phillips A. A review of the performance of tests used to establish whether there is a drug effect in dose-response studies. *Drug Inform J*. 1998;32:683-692.
24. Ruberg SJ. Dose response studies II. Analysis and interpretation. *J Biopharm Stat*. 1995;5:15-42.
25. Sheiner LB, Beal SL, Sambol NC. Study designs for dose-ranging. *Clin Pharmacol Ther*. 1989;46:63-77.
26. Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*. 2005;61:738-748.
27. EMA qualification opinion of MCP-mod as an efficient statistical methodology for model-based design and analysis of phase II dose finding studies under model uncertainty. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2014/02/WC500161027.pdf
28. FDA fit-for-purpose determination of the MCP-mod method for dose finding. 2016. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM508700.pdf
29. Gibson EW. Leadership in statistics: increasing our value and visibility. *Am Stat*. 2019;73:109-116.
30. Bonate PL. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*. Springer; 2011.
31. Lavielle M. *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. Chapman & Hall/CRC; 2014.
32. Gabrielsson J, Weiner D. *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications*. Swedish Pharmaceutical Press; 2016.
33. Marshall S, Burghaus R, et al. Good practices in model-informed drug discovery and development: practice, application, and documentation. *CPT Pharmacomet Syst Pharmacol*. 2016;5:93-122.
34. Marshall S, Madabushi R, Manolis E, et al. Model-informed drug discovery and development: current industry good practice and regulatory expectations and future perspectives. *CPT Pharmacomet Syst Pharmacol*. 2019;8:87-96.
35. FDA Model-Informed Drug Development Pilot Program. 2018. https://www.fda.gov/drugs/development-resources/model-informed-drug-development-pilot-program
36. Cook SF, Bies RR. Disease progression modeling: key concepts and recent developments. *Curr Pharmacol Rep*. 2016;2:221-230.
37. Chen Y, Ni X, Fleisher AS, et al. A simulation study comparing slope model with mixed-model repeated measure to assess cognitive data in clinical trials of Alzheimer's disease. *Alzheimers Dement*. 2018;4:46-53.

38. Romero K, Ito K, Rogers J, et al. The future is now: model-based clinical trial design for Alzheimer's disease. *Clin Pharmacol Ther*. 2015;97:210-214.
39. Upreti VV, Venkatakrishnan K. Model-based meta-analysis: optimizing research, development, and utilization of therapeutics using the totality of evidence. *Clin Pharmacol Ther*. 2019;106:981-992.
40. Demin I, Hamrén B, Luttringer O, Pillai G, Jung T. Longitudinal model-based meta-analysis in rheumatoid arthritis: an application toward model-based drug development. *Clin Pharmacol Ther*. 2012;92:352-359.
41. Leil TA, Lu Y, Bouillon-Pichault M, Wong R, Nowak M. Model-based meta-analysis compares DAS28 rheumatoid arthritis treatment effects and suggests an expedited trial design for early clinical development. *Clin Pharmacol Ther*. 2021;109:517-527.
42. Yoshioka H, Sato H, Hatakeyama H, Hisaka A. Model-based meta-analysis to evaluate optimal doses of direct oral factor Xa inhibitors in atrial fibrillation patients. *Blood Advances*. 2018;2:1066-1075.
43. Teng Z, Gupta N, Hua Z, et al. Model-based meta-analysis for multiple myeloma: a quantitative drug-independent framework for efficient decisions in oncology drug development. *Clin Transl Sci*. 2018;11:218-225.
44. Chan P, Bax L, Chen, C, et al. Model-based meta-analysis on the efficacy of pharmacological treatments for idiopathic pulmonary fibrosis. *CPT Pharmacometrics Syst Pharmacol*. 2017;6:695-704.
45. Sverdlov O, Wong WK, Ryeznik Y. Adaptive clinical trial designs for Phase I cancer studies. *Stat Surveys*. 2014;8:2-44.
46. Collins JM, Grieshaber CK, Chabner BA. Pharmacologically guided phase I clinical trials based upon preclinical drug development. *J Natl Cancer Inst*. 1990;82:1321-1326.
47. Piantadosi S, Liu G. Improved designs for dose escalation studies using pharmacokinetic measurements. *Stat Med*. 1996;15:1605-1618.
48. Patterson S, Francis S, Ireson M, Webber D, Whitehead J. A novel Bayesian decision procedure for early-phase dose-finding studies. *J Biopharm Stat*. 1999;9:583-597.
49. Ursino M, Zohar S, Lentz F, et al. Dose-finding methods for phase I clinical trials using pharmacokinetics in small populations. *Biom J*. 2017;59:804-825.
50. Takeda K, Komatsu K, Morita S. Bayesian dose-finding phase I trial design incorporating pharmacokinetic assessment in the field of oncology. *Pharmaceut Stat*. 2018;17:725-733.
51. Dodds M, Salinger D, Mandema J, Gibbs J, Gibbs M. Clinical trial simulation to inform Phase 2: comparison of concentrated vs. distributed first-in-patient study designs in psoriasis. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:1-9.
52. Sheiner LB, Hashimoto Y, Beal SL. A Simulation study comparing designs for dose ranging. *Stat Med*. 1991;10:303-321.
53. Aoki Y, Röshammar D, Hamrén B, Hooker AC. Model selection and averaging of nonlinear mixed-effect models for robust Phase III dose selection. *J Pharmacokinet Pharmacodyn*. 2017;44:581-597.
54. Buatois S, Ueckert S, Frey N, Retout S, Mentré F. cLRT-Mod: An efficient methodology for pharmacometric model-based analysis of longitudinal Phase II dose finding studies under model uncertainty. *Stat Med*. 2021;40:2435-2451.
55. Sverdlov O, Ryeznik Y, Wong WK. On optimal designs for clinical trials: an updated review. *J Stat Theory Pract*. 2020;14:1-29.
56. Papathanasiou T, Strathe A, Overgaard RV, Lund TM, Hooker AC. Optimizing dose-finding studies for drug combinations based on exposure-response models. *AAPS J*. 2019;21:1-11.
57. King J, Wong WK. Optimal minimax designs for prediction in heteroscedastic models. *J Stat Plan Inference*. 1998;69:371-383.
58. King J, Wong W-K. Minimax D-optimal designs for the logistic model. *Biometrics*. 2000;56:1263-1267.
59. Strömberg EA. Applied adaptive optimal design and novel optimization algorithms for practical use. 2016. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-308452
60. Pierrillas PB, Fouliard S, Chenel, Hooker AC, Friberg LE, Karlsson MO. Model-based adaptive optimal design (MBAOD) Improves combination dose finding designs: an example in oncology. *AAPS J*. 2018;20:39.
61. Chen R-B, Chang S-P, Wang W, Tung H-C, Wong WK. Minimax optimal designs via particle swarm optimization methods. *Stat Comput*. 2015;25:975-988.
62. Masoudi E, Holling H, Wong WK. Application of imperialist competitive algorithm to find minimax and standardized maximin optimal designs. *Comput Stat Data Anal*. 2017;113:330-345.
63. Shi Y, Zhang Z, Wong WK. Particle swarm based algorithms for finding locally and Bayesian D-optimal designs. *J Stat Distribut Appl*. 2019;6:1-17.
64. Wadsworth I, Hampson LV, Jaki T. Extrapolation of efficacy and other data to support the development of new medicines for children: a systematic review of methods. *Stat Methods Med Res*. 2018;27:398-413.
65. Wadsworth I, Hampson LV, Bornkamp B, Jaki T. Exposure-response modelling approaches for determining optimal dosing rules in children. *Stat Methods Med Res*. 2020;29:2583-2602.
66. Leong R, Vieira MLT, Zhao P, et al. Regulatory experience with physiologically based pharmacokinetic modeling for pediatric drug trials. *Clin Pharmacol Ther*. 2012;91:926-931.
67. Turner MA, Catapano M, Hirschfeld S, Giaquinto C. Paediatric drug development: the impact of evolving regulations. *Adv Drug Deliv Rev*. 2014;73:2-13.
68. Jadhav PR, Zhang J, Gobburu JVS. Leveraging prior quantitative knowledge in guiding pediatric drug development: a case study. *Pharmaceut Stat*. 2009;8:216-224.
69. Jadhav PR, Kern SE. The need for modeling and simulation to design clinical investigations in children. *J Clin Pharmacol*. 2010;50:121S-129S.
70. Gamalo-Siebers M, Savic J, Basu C, et al. Statistical modeling for Bayesian extrapolation of adult clinical trial information in pediatric drug evaluation. *Pharmaceut Stat*. 2017;16:232-249.
71. Almquist J, Sadiq MW, Eriksson UG, Myrbäck TH, Prothon S, Leander J. Estimation of equipotent doses for anti-inflammatory effects of prednisolone and AZD9567, an oral selective nonsteroidal glucocorticoid receptor modulator. *CPT Pharmacomet Syst Pharmacol*. 2020;9:444-455.
72. FDA Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products. 2019. https://www.fda.gov/media/121320/download
73. Kosorok MR, Laber EB. Precision medicine. *Ann Rev Stat Its Appl*. 2019;6:263-286.
74. Lai TL, Lavori PW, Tsang KW. Adaptive enrichment designs for confirmatory trials. *Stat Med*. 2019;38:613-624.
75. Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med*. 2017;377:62-70.

76. Nair S, Kong A-NT. Emerging roles for clinical pharmacometrics in cancer precision medicine. *Cur Pharmacol Rep*. 2018;4:276-283.
77. Mould DR, Lesko LJ. Personalized medicine: integrating individual exposure and response information at the bedside. In: Schmidt, S. & Derendorf, H., eds. *Applied Pharmacometrics*. AAPS Press Springer; 2014: 65-82.
78. FDA CDRH approval letter for the 670G system (PMA Applicant: Medtronic MiniMed, Inc.). 2016. https://www.accessdata.fda.gov/cdrh_docs/pdf16/P160017a.pdf
79. Karlsson K, Vong C, Bergstrand M, Jonsson E, Karlsson M. Comparisons of analysis methods for proof-of-concept trials. *CPT Pharmacomet Syst Pharmacol*. 2013;2:1-8.
80. Jones AK, Salem AH, Freise KJ. Power determination during drug development: is optimizing the sample size based on exposure-response Analyses underutilized? *CPT Pharmacomet Syst Pharmacol*. 2019;8:138-145.
81. A roadmap to using randomization in clinical trials. *BMC Med Res Methodol*. https://doi.org/10.21203/rs.3.rs-135735/v1
82. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc*. 1952;47:583-621.
83. Wang Y, Rosenberger WF, Uschner D. Randomization tests for multiarmed randomized clinical trials. *Stat Med*. 2020;39:494-509.
84. Wählby U, Jonsson EN, Karlsson MO. Assessment of actual significance levels for covariate effects in NONMEM. *J Pharmacokinet Pharmacodyn*. 2001;28:231-252.
85. Deng C, Plan EL, Karlsson MO. Influence of clinical trial design to detect drug effect in systems with within subject variability. 2015;24. www.page-meeting.org/?abstract=3549
86. Keizer R, Karlsson M, Hooker A. Modeling and simulation workbench for NONMEM: tutorial on Pirana, PsN, and Xpose. *CPT Pharmacomet Syst Pharmacol*. 2013;2:1-9.
87. Rosenkranz GK. The impact of randomization on the analysis of clinical trials. *Stat Med*. 2011;30:3475-3487.
88. Sverdlov O, Ryeznik Y. Implementing unequal randomization in clinical trials with heterogeneous treatment costs. *Stat Med*. 2019;38:2905-2927.
89. Zhang L, Sinha V, Forgue ST, et al. Model-based drug development: the road to quantitative pharmacology. *J Pharmacokinet Pharmacodyn*. 2006;33:369-393.
90. Nayak S, Sander O, Al-Huniti N, et al. Getting innovative therapies faster to patients at the right dose: impact of quantitative pharmacology towards first registration and expanding therapeutic use. *Clin Pharmacol Ther*. 2018;103:378-383.
91. Wang Y, Jadhav PR, Lala M, Gobburu JV. Clarification on precision criteria to derive sample size when designing pediatric pharmacokinetic studies. *J Clin Pharmacol*. 2012;52:1601-1606.
92. Salem F, Ogungbenro K, Vajjah P, Johnson TN, Aarons L, Rostami-Hodjegan A. Precision criteria to derive sample size when designing pediatric pharmacokinetic studies: which measure of variability should be used? *J Clin Pharmacol*. 2014;54:311-317.
93. Maloney A, Karlsson MO, Simonsson USH. Optimal adaptive design in clinical drug development: a simulation example. *J Clin Pharmacol*. 2007;47:1231-1243.
94. Foo LK, Duffull S. Adaptive optimal design for bridging studies with an application to population pharmacokinetic studies. *Pharm Res*. 2012;29:1530-1543.
95. Germovsek E, Barker CIS, Sharland M, Standing JF. Scaling clearance in paediatric pharmacokinetics: all models are wrong, which are useful? *Br J Clin Pharmacol*. 2017;83:777-790.
96. Venkatakrishnan K, Cook J. Driving access to medicines with a totality of evidence mindset: an opportunity for clinical pharmacology. *Clin Pharmacol Ther*. 2018;103:373-375.
97. Wagner JA, Dahlem AM, Hudson LD, et al. Application of a dynamic map for learning, communicating, navigating, and improving therapeutic development. *Clin Transl Sci*. 2018;11:166-174.
98. Venkatakrishnan K, Zheng S, Musante CJ, et al. Toward progress in quantitative translational medicine: a call to action. *Clin Pharmacol Ther*. 2020;107:85-88.
99. Koch G, Pfister M, Daunhawer I, et al. Pharmacometrics and machine learning partner to advance clinical data analysis. *Clin Pharmacol Ther*. 2020;107:926-933.
100. Sheiner LB. The intellectual health of clinical drug evaluation. *Clin Pharmacol Ther*. 1991;50:4-9.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ryeznik Y, Sverdlov O, Svensson EM, Montepiedra G, Hooker AC, Wong WK. Pharmacometrics meets statistics—A synergy for modern drug development. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:1134–1149. https://doi.org/10.1002/psp4.12696