Citation: CPT Pharmacometrics Syst. Pharmacol. (2017) 6, 418–429; doi:10.1002/psp4.12196

WHITE PAPER

Advanced Methods for Dose and Regimen Finding During Drug Development: Summary of the EMA/EFPIA Workshop on Dose Finding (London 4–5 December 2014)

FT Musuamba[1,2,3]*, E Manolis[1,4], N Holford[5], SYA Cheung[6], LE Friberg[7], K Ogungbenro[8], M Posch[9], JWT Yates[6], S Berry[10], N Thomas[11], S Corriol-Rohou[6], B Bornkamp[12], F Bretz[9,12], AC Hooker[7], PH Van der Graaf[13,14], JF Standing[1,15], J Hay[1,16], S Cole[1,16], V Gigante[1,17], K Karlsson[1,18], T Dumortier[12], N Benda[1,19], F Serone[1,17], S Das[6], A Brochot[20], F Ehmann[4], R Hemmings[16] and I Skottheim Rusten[1,21]

Inadequate dose selection for confirmatory trials is currently still one of the most challenging issues in drug development, as illustrated by high rates of late-stage attritions in clinical development and postmarketing commitments required by regulatory institutions. In an effort to shift the current paradigm in dose and regimen selection and highlight the availability and usefulness of well-established and regulatory-acceptable methods, the European Medicines Agency (EMA) in collaboration with the European Federation of Pharmaceutical Industries Association (EFPIA) hosted a multistakeholder workshop on dose finding (London 4–5 December 2014). Some methodologies that could constitute a toolkit for drug developers and regulators were presented. These methods are described in the present report: they include five advanced methods for data analysis (empirical regression models, pharmacometrics models, quantitative systems pharmacology models, MCP-Mod, and model averaging) and three methods for study design optimization (Fisher information matrix (FIM)-based methods, clinical trial simulations, and adaptive studies). Pairwise comparisons were also discussed during the workshop; however, mostly for historical reasons. This paper discusses the added value and limitations of these methods as well as challenges for their implementation. Some applications in different therapeutic areas are also summarized, in line with the discussions at the workshop. There was agreement at the workshop on the fact that selection of dose for phase III is an estimation problem and should not be addressed via hypothesis testing. Dose selection for phase III trials should be informed by well-designed dose-finding studies; however, the specific choice of method(s) will depend on several aspects and it is not possible to recommend a generalized decision tree. There are many valuable methods available, the methods are not mutually exclusive, and they should be used in conjunction to ensure a scientifically rigorous understanding of the dosing rationale.
CPT Pharmacometrics Syst. Pharmacol. (2017) **6**, 418–429; doi:10.1002/psp4.12196; published online 0 Month 2017.

In order to receive marketing authorization (MA) and to be maintained on the market, new drug candidates need to demonstrate good evidence of efficacy and safety in the sought indication. Adequately powered phase III randomized controlled trials (RCTs) where the new drug candidate is tested against placebo or an active comparator is the gold standard for confirmation of efficacy and safety. Although it is logical to assume that medicinal products that advance to phase III trials are adequately characterized in terms of pharmacokinetics (PK), pharmacodynamics (PD), and the efficacy and safety profile in earlier stages of drug development, the high attrition rate in phase III does not support this.[1] One of the contributing factors to this high attrition rate is inadequate dose and regimen selection and, more generally, the insufficient understanding of the pharmacology to design an optimal phase III program.[2–4] Even successful phase III trials and regulatory labeling may not include the optimal dose and regimen, especially for special populations such as the elderly and pediatrics, as shown in postmarketing commitments (Post Authorisation Efficacy Studies, (PAES); Post Authorisation Safety Studies (PASS)) and through subsequent changes to the dosing recommendations postmarketing.[5]

The most commonly used method for defining the dosing rationale (dose and dosing regimen) is the pairwise comparison of different doses with a common control (e.g., placebo), a method based on minimal assumptions which, however, has known limitations in this context, including reliance on *P*-values and the need for the dose ranging studies to be powered for multiple comparisons.[6] Further, the exploratory development is often poorly conceived by focusing on selecting the dose/regimen instead of

[1]EMA Modelling and Simulation Working Group, London, UK; [2]Federal Agency for Medicines and Health Products, Brussels, Belgium; [3]UMR850 INSERM, Université de Limoges, Limoges, France; [4]European Medicines Agency, London, UK; [5]Department of Pharmacology & Clinical Pharmacology, University of Auckland, Auckland, New Zealand; [6]AstraZeneca UK Limited, London, UK; [7]Uppsala University, Uppsala, Sweden; [8]Manchester University, Manchester, UK; [9]Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Vienna, Austria; [10]Berry consultants, Austin, Texas, USA; [11]Pfizer, London, UK; [12]Novartis, London, UK; [13]Leiden Academic Centre for Drug Research, Leiden, The Netherlands; [14]Certara QSP, Canterbury, UK; [15]University College London, London, UK; [16]Medicines and Healthcare Products Regulatory Agency, London, UK; [17]Agenzia Italiana del Farmaco, Roma, Italy; [18]Medical Products Agency, Uppsala, Sweden; [19]Bundesinstitut für Arzneimittel und Medizinprodukte, Bonn, Germany; [20]ABLYNX, Gent, Belgium; [21]Norvegian Medicines Agency, Oslo, Norway. *Correspondence: F Musuamba (Flora.MusuambaTshinanu@fagg-afmps.be)
Received 21 November 2016; accepted 27 March 2017; published online on 0 Month 2017. doi:10.1002/psp4.12196

characterizing dose–exposure–response (DER). The latter should be prioritized, since it enables more efficient dose/regimen selection and even goes beyond that. DER can be used to inform summary of product characteristics (SmPC), risk management plan (RMP), and support extrapolation to different populations. In addition, comprehensive DER information may be used as a sole basis for approval of a different dose/regimen, and a new formulation/route of administration postmarketing.

This body of evidence should be ideally collected ahead of confirmatory trials in order to be prospectively and adequately tested.[7] In a well-designed development program, pivotal trials should not enable characterization of DER because all the patients included should show optimal response. Inevitably, though, as the drug development is focused initially on a well-characterized population, which expands later to reflect the target population in subsequent stages, DER findings may occur. It is expected that DER in phase III will be a signal detection exercise that will evaluate consistency of drug effects across subgroups and propose further actions if inconsistencies are shown.

In an effort to shift the current paradigm in dose and regimen selection and highlight the availability and usefulness of well-established and regulatory-acceptable tools and methods, the European Medicines Agency (EMA) in collaboration with the European Federation of Pharmaceutical Industries Association (EFPIA) hosted a multistakeholder workshop on dose finding (London 4–5 December 2014). Some methodologies that could constitute a toolkit for drug developers and regulators were presented. This paper will discuss their added values, limitations, and challenges. A separate paper will discuss how regulators value information on dose–exposure–response relationships at the stage of MAA and postmarketing.

Building a scientifically rigorous justification for the dosing rationale for a phase III trial will depend on a large number of factors that could make this exercise rather case-specific. It may, however, be possible to benchmark the different tools available and propose some guiding principles on when and how to use them.

Methods presented during the workshop can be divided into two categories: 1) methods for data analysis and 2) methods for design optimization.

The aim of this paper is to convey the message that dose/regimen finding and selection should be a multidisciplinary team decision and that there are many methods available to make sufficiently informed decisions rather than to propose a general decision tree applicable for all drug developments in different therapeutic areas. This will be illustrated through applications where some advanced methodologies have been applied with relatively positive outcome.

CHALLENGES FOR DOSE SELECTION DURING DRUG DEVELOPMENT

The traditional drug development is usually concerned with what dose should be chosen for phase III, to meet the requirements for prespecification and strict type I error control in confirmatory trials. Even if this paradigm facilitates confirmatory testing, at the end it is not clear if the selected dose is optimal because the DER space is not known. In addition, the confirmatory data generated in phase III with the specific dose makes it difficult to challenge "the dose." There is an agreement in different fora, including the EMA dose-finding workshop, that there is a way to improve dose selection for phase III, retaining the principles of confirmatory testing.

One principle-based method of dose finding in drug development that can also be linked directly to the clinical use of a drug was described more than 20 years ago.[8] The principle is based on the concept of a target effect and use of a PD model to predict the target concentration needed to achieve that effect and a PK model to predict the dose required to achieve the target concentration (Eqs. 1–4).

$$\textit{Target Effect} \rightarrow \textit{Target Concentration} \rightarrow \textit{Dose} \quad (1)$$

$$\textit{Target Concentration} = C50 \times \textit{Target Effect} / (\textit{E}\max - \textit{Target Effect}) \quad (2)$$

$$\textit{Loading Dose} = \textit{Volume of Distribution} \times \textit{Target Concentration} \quad (3)$$

$$\textit{Maintenance Dose Rate} = \textit{Clearance} \times \textit{Target Concentration} \quad (4)$$

This simple example shows that only four parameters are required to find the right dose. These are maximum drug effect (E-max) and potency (C50) to predict the target concentration and volume of distribution (V) and clearance (CL) to predict the dose. In actual drug development programs, the PD and PK models are usually more complex but the principle remains the same.

The challenges of applying this model-based method for dose finding are:

1. What is the target effect? From the patient perspective the target effect is to relieve symptoms completely or restore normal function or return to a normal survival probability. In practice, this is rarely achievable because drugs may not be sufficiently effective to achieve these goals and/or the magnitude of the effect is limited by adverse effects. The desired beneficial effects and the undesired adverse effects are both determined by drug concentration. The challenge is to thus determine the optimal target concentration that achieves as much benefit as possible with acceptable adverse effects. The challenge is made more complex because the time course of beneficial and adverse effects may be different, so that the target concentration itself may change with time. This is the implicit principle behind the empirical dosing aphorism "start low and go slow" (e.g., Ref. 9).
2. What is the concentration (or related exposure metric) that produces the effect? This is a more general problem than simply predicting the target concentration because it must be clearly understood what exposure metric is most appropriate for a target. When drugs have reversible effects and the clinically useful concentrations are close to and above the C50 and the desired clinical benefit is dependent on a cumulative drug effect (e.g., frusemide for relief of heart failure symptoms), then schedule dependence is expected and the time course of concentration must be considered. For drugs which have slowly reversible actions or are irreversible and require a new drug target to be synthesized, then the cumulative

concentration over time ("area under the curve" or AUC) and thus the cumulative dose is the target (e.g., busulfan for bone marrow ablation). In between these extremes of the full time course to the full integrated time course the most widely used concentration metric is the average steady state concentration (Css). Css is directly related to the steady state AUC in a dosing interval. AUC is commonly recommended but it requires that the dosing interval be defined as well, which can lead to confusion when this is not clearly stated. Css avoids this problem and is simple to use to predict the required maintenance dose. Specific target concentration (or related metrics) can be evaluated by concentration-controlled clinical trials.[10–14] Further challenges arise when dealing with drug combination therapy but the principle is the same.

3. What variability of effect is acceptable when dosing is based on a population target concentration? How does it translate into what variability is acceptable in the exposure metric(s)? Population analysis methods applied to PK and PD models can describe the predictable and unpredictable (random) components of variability, but the challenge is identifying the acceptable range that allows the drug to be used safely and effectively.[15]
4. Finally, some practical considerations are also of importance. The scientific challenges in defining the target effect, concentration, and finally dose are further confounded by the limitations of clinical trials, i.e., restricted number of events, patients, and investigations that can be conducted. For example, in case of lifelong treatments to prevent rare events, the DER relationship is often built on PD markers, with the additional complexity of translation to clinical outcome. In rare diseases, where the number of patients is limited, there is a tradeoff between the patients enrolled in dose selection vs. the ones "reserved" for confirmatory testing. In some diseases, such as tuberculosis, it is challenging to measure concentrations at the site of infection, e.g., lung granuloma; therefore, assumptions are made on the distribution of the medicinal products in the lungs and actual PK measurements are limited to plasma and epithelial lining fluid.

METHODS

The different methods presented during the workshop for data analysis and/or study design optimization (**Figure 1**) are described in detail in the following sections.

Methods for data analysis

Pairwise comparisons. Pairwise comparisons were not thoroughly discussed during the workshop, given their limitations. This method will therefore only be shortly presented in this paper. However, given that it has historically been the main method used and that it is still widely used for dose selection in phase II studies, several illustrative examples can be found in the literature.[16,17] Generally, parallel-group, double-blind, randomized, dose-ranging trials are implemented, and patients are randomly assigned to one of the dose levels and/or placebo/active comparator. The aim of the analysis is often the selection of the adequate dose for confirmatory trials. The results obtained in different treatment arms are compared with a statistical test (e.g., analysis of variance (ANOVA)). This method, while allowing to make a statement on specific dose without relying on model assumptions, penalizes the developer when a wide range of doses are tested, has limited statistical power compared to model-informed methods, and makes inefficient use or disregards the available information.[17]

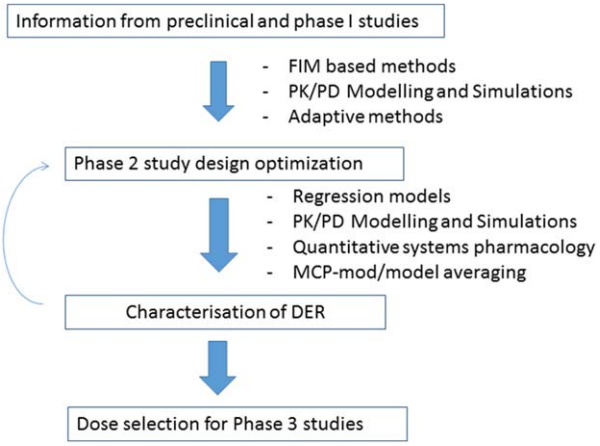

Figure 1 Schematic representation of the role of advanced methods in phase II and III drug development.

During the workshop it was agreed that selection of dose for phase III is an estimation problem and should not be addressed via hypothesis testing.

Advanced methods

Empirical dose (exposure) response models (regression models). Empirical regression methods (such as linear, curvilinear spline, and E-max models, etc.) are extensively used to describe phase II dose-ranging results. They aim to (empirically) compare and analyze the dose–response relationship based on different doses tested in phase II trials and use this as a basis to select a dose (usually one of the tested doses).[18,19] Most of these methods are optimal under certain conditions; however, many of the statistical methods have been optimized for conditions seldom encountered in practice. Two meta-analyses have recently been conducted on dose–response studies from small-molecule drugs, which aimed to describe the designs of the studies, and to summarize methods used for dose–response analysis. The results showed that, for the majority of cases, E-max models were used for dose selection.[18,19]

The E-max model is very common in pharmacology, and can be derived from the law of mass action relating to receptor occupancy.[20] Other similar families of models monotonically increasing (or decreasing) to an asymptote, such as models formed by rescaling the logistic distribution, could also be used. When unusual situations are encountered, they are often more appropriately assessed by consideration of other approaches such as better PK characterization, description of reasons for discontinuation, or more mechanistic (physiologically based) PK and/or PK/PD models.

Multiple comparison procedures and modeling (MCP-Mod). MCP-Mod is a principled method for empirical dose–response analysis. The modeling is based on a set of candidate dose–response shapes that is predefined at the design stage of the trial. The methodology consists of two

steps: (i) A statistical test for a dose–response signal, assessing whether there is a dose-related trend over placebo (the MCP step). In the second step (ii) the dose–response curve is estimated based on model selection or model averaging techniques. For more details on MCP-Mod, we refer to Bretz *et al*. and Pinheiro *et al*.[21,22] MCP-Mod was qualified as an efficient statistical methodology for model-based design and analysis of phase II dose finding studies under model uncertainty by the CHMP,[23] and received a fit-for-purpose determination by the US Food and Drug Administration (FDA).[24]

MCP-Mod applies when the primary focus of a phase II study is the investigation of the population dose–response curve at a specific timepoint. The method is most informative if it is applied at a timepoint when the dose–response achieves steady state. It applies to both parallel groups and crossover designs and when at least three active doses are utilized vs. a placebo (or placebo-like) comparator. The method does not apply if the major focus of the study is to determine an adequate administration frequency rather than the dose itself within given regimens. It is possible to extend MCP-Mod to allow modeling more than one administration frequency when making certain assumptions, e.g., some of the model parameters are shared between regimens. MCP-Mod also does not apply when very different irregular treatment regimens are compared (e.g., regimens with and without loading dose), or different administration forms are compared. While the original approach was not designed to model the evolvement over time, extensions have been described based on longitudinal nonlinear mixed-effects models.[25] The method should not be used when patients are titrated based on efficacy and/or safety outcomes in the design, because a population dose–response analysis is misleading in this situation. The method also does not apply to safety dose escalation studies.

It has been demonstrated through simulations in a number of realistic scenarios that MCP-Mod is an improvement over pairwise testing approaches, in terms of power to detect a dose–response signal and to estimate the dose–response curve (see Bornkamp *et al*., among others).[26,27] The method is intended to be specified at the design stage. This means that the study design, including the total sample size and the dose levels, should be selected to reliably determine the dose–response curve. This encourages more informative study designs. Such a structured modeling approach makes the modeling process more transparent and less subject to cherry-picking or overfitting. MCP-Mod can be implemented relatively quickly, compared to a longitudinal DER analysis. The caveat of prespecification (e.g., for relevant covariates, times of measures, etc.) is that it might not be flexible enough to accommodate trends in the data unanticipated at the design stage. It is therefore recommended to perform additional exploratory *post-hoc* analyses (not necessarily based on MCP-Mod) evaluating the effect of covariates on different aspects of the dose–response curve; these analyses should be interpreted carefully. In its standard form, MCP-Mod can only include covariates as additional effects to the intercept of the dose–response model, and covariates are only included if they are assumed to affect the response at the design stage. Also, MCP-Mod does not use exposure data, but cross-sectional exposure–response analyses could be performed with MCP-Mod as well. A potential advantage of MCP-Mod as compared to single models is that model uncertainty in the estimation (e.g., of the target dose or the dose–response curve) can be addressed using appropriate techniques. This solves the impact of model selection uncertainty on the final inferences to a large extent.

Model averaging. If a nonlinear regression model is used to quantify the dose–response (D-R) relationship (or DER relationship), a natural question is which model to use. In general, even with high amounts of prior information, there will be uncertainty in terms of model structure in addition to model parameter values when fitting and interpreting DER data. Moreover, while several models might fit the data well, they may differ on certain estimated quantities of interest; for example, the target dose estimate as well as the uncertainty around that estimate. There have been several attempts to weaken assumptions about the choice of the model to use (model building) through model selection and model averaging of predefined candidate models of interest.[28,29]

Model averaging acknowledges model uncertainty explicitly as part of the inference and "averages" over all the model candidates using predefined relative weights for each model. Averaging of model candidates can be performed by using either a frequentist or a Bayesian approach. In each approach, there are methods for model weighing and for averaging the outcome (e.g., the smallest dose producing an effect that is at least as large as the target effect). When a frequentist approach is used, the likelihood of the parameter estimates can be approximated and then different information criteria (Akaike information criterion (AIC), Bayesian information criterion (BIC), etc.) can be used as selection criteria or as the weights in an averaging scheme. These effects from each model should be evaluated in such a way as to account for parameter uncertainty. Other methods of model averaging that utilize bootstrapped model selection techniques to average over model predictions have also shown promise.[30–34]

Model averaging techniques can increase the probability of making the correct dose-finding decisions compared to a single model informed analysis or to conventional nonmodel informed study protocols. There exists empirical evidence that model averaging also improves the estimation efficiency. As a consequence of including model structure uncertainty and parameter uncertainty when using model averaging, the quantified uncertainty of an effect may appear to be bigger than with a single model informed prediction. However, single model informed predictions may be overoptimistic or biased (if the "wrong" model is used). As more examples of the dangers of ignoring model uncertainty are publicized, as computing power continues to expand, and as the size of databases, the numbers of variables, and hence the numbers of possible models increase, accounting for model uncertainty (e.g., application through model sensitivity analysis) will become an integral part of statistical modeling.[28,29,34,35]

The challenges described for MCP-mod related to model prespecification are also applicable to model averaging.

Pharmacometrics models. The traditional use of pharmacometric (PMx) modeling for dose selection originates from the concept of "learn and confirm," which emerged in the late 1990s and emphasized the need for early development activities to more effectively inform later-stage development.[36,37] The use of modeling in drug development has increased considerably in the last two decades; models are used to describe PK and PD properties of drugs and have now become a standard tool in the pharmaceutical industry.[38–41] While PK models relate observed concentrations in biological fluids to the times of measurements and the administered dose, (PK/)PD models relate observed response (clinical efficacy or toxicity) or their biomarkers/surrogates to dose or concentrations.

Mixed-effects models (population approach) have been widely used for the analysis of PK, PD, and PK/PD data. This allows estimation of population parameters that describe the average response and a measure of variability in the data. The importance of using this approach during drug development has been widely discussed in the literature; it is particularly attractive for unbalanced and sparse data.[42–44]

In dose-finding studies, the response of interest can be continuous or discrete. Continuous response in clinical trials can take any value and examples include blood pressure, heart rate, tumor size, time-course of leukocytes, and neutrophils and Forced Expiratory Volume in 1 second (FEV1).[45,46] Discrete responses, on the other hand, are responses that can only take a finite number of values, they can be ordinal, nominal, count, or time-to-event variables. Examples of discrete variable include seizure count in epilepsy, grade of adverse events, and survival data in oncology.[47–50] Dose optimization for efficacy and safety can sometimes involve a mixture of continuous and discrete outcomes in a multiresponse model where a careful balance of the outcomes is important.[51]

A number of structural models such as linear, E-max, exponential, log-linear, logistic models, and indirect response models have been used to characterize either dose– or exposure–response profiles as part of PMx models.[52,53]

As far as the link to physiology and pharmacology are concerned, different types of PMx models have been described including empirical and (semi)mechanistic models.[54–56] Model-based meta-analysis (MBMA) has also been proposed when different (competing) models are available to describe the data.[57]

Quantitative and systems pharmacology models. Quantitative and systems pharmacology (QSP) models are mechanistic disease progression, PK/PD, and physiologically based (PB) PK models that focus on describing in a quantitative manner the interactions between the (healthy and diseased) organism (so-called system) and different drugs or drug candidates, starting from the characterization and quantification of a network of biological/molecular mechanistic pathways. When QSP methods are used for dose selection, there is therefore a clear need for understanding, in a precise and predictive manner, how drugs modulate the system. Both temporal and spatial scales are considered for this purpose. They incorporate data at several temporal and spatial scales and consider interactions among all relevant elements at different levels (biomolecules, cells, tissues, organs, and organism). These methods have also been widely described in the literature.[58–62]

QSP models constitute powerful tools to quantitatively understand and predict therapeutic and toxic effects of drugs. QSP models, in various therapeutic areas, have successfully been used to inform optimal dose selection during the exploratory phase of drug development.[58–62]

With regard to characterization of drug efficacy, QSP models include investigation of the sources of variability in drug response at different levels: single-cell, organ, and/or patient level. Variability that can originate from intrinsic or environment factors. QSP approaches include incorporation of data from diverse sources (from omics to clinical levels), characterization of pharmacodynamic biomarkers, development of inform integrated and multiscale models of drug-response in different patient populations, and development of animal and tissue models for preclinical pharmacology. The final aim of using a QSP approach for efficacious dose finding is better target validation, development of multiscale pharmacological models, and therefore fewer phase II and III failures due to lack of efficacy.[58]

QSP models have also been used for drug safety prediction and evaluation in a quantitative manner, and therefore offer the possibility to have quantitative metrics on both sides of the benefit/risk (B/R) balance so that statements can be made about clinical benefit, clinical utility, or effectiveness of drugs under development.[62]

Phase II studies are particularly relevant to systems pharmacology because it is at this stage that correct selection of targets, drugs, dosing regimens, and therapeutic effects should rigorously be tested.

It should, however; be noted that the actual mechanism of action (i.e., how target engagement leads to patient responses) remains at least partly obscure for most therapeutic drugs. An immediate challenge for use of QSP will therefore be developing the capacity to make rigorous statements about both sensitivity and specificity for predicted treatment outcomes.[58]

Moreover, use of QSP will require the participation of investigators with different skill sets, but will also benefit from the insight of individuals who can bring concepts and information from several relevant fields (integrators rather than specialists).[58]

This poses a challenge, since the current culture of academic biomedicine research community has evolved to reward individuals who are embedded in the culture of a single field, with the unintended consequence that interdisciplinary innovation is made harder.[58]

Methods for study design optimization

Irrespective of the method used for the analysis of data, dose selection for confirmatory trials will also be impacted by design factors in phase II dose-ranging studies and vice versa. Design factors include dose range, study durations,

number of patients included (study size), timing, and number of PK/PD/clinical response measures, etc. All the inclusion/exclusion criteria can also be considered as part of study design and will impact the results of phase II data analysis and subsequent phase III dose selection. Effective tools for study design optimization are described in the literature. Several methods have been described for this purpose, including optimality defined based on the determinant of the Fisher information matrix (FIM) and clinical trial simulations. These methods should be used when planning phase II studies. Design optimization tools work better if the model to be used to analyze phase II data has been characterized from previous experience/study(ies). Otherwise, a two-stage approach should be envisaged in which the study design implemented at the initiation of the study is refined based on the results of the interim analysis. Three methods for phase II study design optimization are detailed in the following sections: FIM-based methods, clinical trial simulations, and adaptive study designs.

Fisher information matrix-based methods. The use of a model-informed approach in drug development provides a framework that allows studies such as dose-finding studies to be performed in a way that data collection is carried out in a very efficient way for proper characterization of dose–response or exposure–response relationships. Optimal designs are obtained by optimization of a function of the FIM, an approach that is based on Cramer–Rao inequality. Expressions for FIM for fixed and mixed effects model have been described. This is usually a square matrix of the dimension of the number of parameters in the model, each row/column of the matrix corresponding to a parameter. Different optimality criteria (A, C, D, E, G) have been described in the literature, depending on the objective of the study, a function of which can be optimized to obtain optimal doses for characterization of dose–response relationships as uniresponse or multiresponse models with a mixture of both continuous and discrete outcomes. In addition to (number of) dose levels, other design factors can be optimized to ensure collection of informative data in exploratory trials, such as number of patients (study size), number of samples, sampling times, and study duration.[63–69]

A major challenge in optimal design of nonlinear models is that the FIM is a function of unknown parameters. Such parameter estimates can sometimes be obtained from previous studies or another relevant study i.e., another drug in the class. Local validity of the results is another problem; the result of the optimization is only valid for the parameter and model that were used for the optimization. If a set of parameter estimates have been used for the optimization, the results are only valid for this set and cannot be generalized. To address this problem, a robust design using Bayesian approaches can be employed; this allows incorporation of a measure of uncertainty in the parameter values in the optimization. Similarly, methodologies to guard against possible model misspecification when there is more than one competing model have also been described.[70,71]

Clinical trial simulations (CTS). Clinical trial simulation involves the specification of several types of models which are then used together to simulate the outcome of individual patients in a clinical trial to produce a dataset similar to the real data from an actual trial.

These key models are:

1. A model describing structural features and variability of the disease and its progression and the time course of drug effects on the disease.
2. A model for the covariate distribution in the target population that can be used to predict individual efficacy and safety differences that are dependent on the covariates (for known covariates that are associated with the efficacy or safety outcomes).
3. A model for the nominal design of the study involving the treatment arms and dosing strategies.
4. A model for deviations from the nominal design due to trial participant withdrawal, incomplete adherence with the dosing regimen, and missing observations.

These models may be implemented using general-purpose modeling and simulation tools. The ability of clinical trial simulation to simulate many replicates of trial participants and then replicates of trials allow statistical analysis of the simulation results, which in turn can answer questions about the bias and uncertainty of model parameters and the power of hypothesis tests used to evaluate the success of the trial.

Clinical trial simulation builds on optimal design methods used to obtain a nominal design model (Model 3 in the list above). It allows accounting for impact of missing data on the outcome of a trial (Model 4 in the list above). It also allows key uncertainties about components in any of the four models above to be explored as specific scenarios to understand the importance of assumptions in determining the trial analysis outcome. CTS may also be seen as a powerful tool to visualize and examine complex statistical distributions.

As for other model-based methods, the technical aspects of implementing a clinical trial simulation usually involve just one or two experts familiar with the modeling and simulation methodology. But they rely on working with a larger team of clinicians, statisticians, and project managers who are responsible for defining the goals of the clinical trial simulation in order to ensure the results can be used to make decisions about actual clinical trial design.

When this larger team works effectively, it leads to many benefits for the drug development process that go beyond the design of one particular trial. The disease and drug model (Model 1 in the list above) has to be built based on existing knowledge, but commonly important elements of the model needed to simulate a clinical trial are not known. Plausible assumptions about missing components must be agreed on by the wider team before they are tested as simulation scenarios. The process of understanding what is missing often leads to a better appreciation of the overall project plan and better decision-making about how to approach the development of a drug or series of drugs.

There are many examples of successful and not-so-successful clinical trial simulation projects. The use of

clinical trial simulation has been extensively reviewed and discussed, e.g., in Refs. 72–87. The key challenge for taking advantage of clinical trial simulation is for project management to start developing a clinical trial simulation strategy earlier in the drug development process. In many cases clinical trial simulation is only thought of late in the development of a drug when hard choices have to be made about trial design.

Adaptive and Bayesian adaptive studies. In adaptive studies, the study is designed to have flexible aspects that dynamically change during the course of the trial, e.g., based on predetermined algorithms and models that are learning from the accruing data in the trial. Clinical trials are therefore prospectively designed with *adaptive* features to make the studies more efficient (e.g., shorter duration, fewer patients), more likely to demonstrate an effect of the drug if one exists, or more informative about the most therapeutic dose or the go/no-go decision (e.g., by providing broader dose–response information).[88–92] The great potential of adaptive designs is that it allows the trial to learn from the accruing data in order to continue to collect relevant and informative data as the trial continues. For example, in a dose-finding trial a predetermined sample size of 200 can be specified or a flexible sample size between 100 and 300 can be specified, allowing the trial to stop when the appropriate questions are answered. A critical aspect of dose-finding studies is the ability to select an optimal dose. In a fixed design the doses and the allocation ratio of each is prespecified. In an adaptive dose-finding design an algorithm that allows the allocation to vary (response adaptive randomization) during the course of the trial can create more patients being allocated to the more therapeutically beneficial and more informative doses, allowing the potential for better identification of the optimal therapeutic dose with fewer resources. Frequently, response adaptive designs can allow more doses to be explored in dose-finding trials for the same number of patients.

There is a synergy with adaptive trials to the use of Bayesian methods. Many dose-finding trials will have many interim analyses to keep the trial design on track. Bayesian methods work well in this setting because the calculation of the posterior distribution at an interim does not only depend on the result or presence of previous looks in the trial. Frequentist techniques depend on the sample space of the data, and thus can be quite difficult computationally within an adaptive trial design. Additionally, the use of posterior distributions, utility functions, and predictive probabilities make the Bayesian approach powerful within adaptive designs.

The complexity of adaptive designs demands more time for the planning and design of these trials. The flexibility of the designs, driven by the adaptations and models, create a situation where the operating characteristics of the design cannot be easily calculated with analytical methods. Therefore, clinical trial simulation becomes critical to understand the characteristics of, and to optimize the behavior of, the adaptive design. Being "adaptive" does not guarantee a better-performing trial; hence, appropriate exploration and optimization of the adaptive design is critical to having a well-performing design.

APPLICATIONS

Some applications of the above-described methods were presented during the workshop. Therapeutic areas for which examples were presented during the workshop included central nervous system (CNS), oncology, antibiotics, antivirals, cardiovascular, and immunology. For a matter of conciseness, only applications in three therapeutic areas (oncology, infection disease, and organ transplantation) will be briefly discussed. The reader is referred to the workshop summary report[87] for more descriptions of additional applications as discussed during the workshop. A session of the workshop was dedicated to dosing in special populations, which is also summarized in this section.

Oncology drugs

Identification of the maximum tolerated dose (MTD) is still the most commonly used method to identify the recommended phase II dose (RP2D) for oncology compounds. There is a need to reconsider the assessment of MTD for some medicinal products for which continuous dosing is the foreseen schedule, as exemplified by the case of the multitargeted tyrosine kinase inhibitor (TKI, cabozantinib), a conventional $3 + 3$ design was used to identify MTD. The need for dose reduction in a high percentage of patients was confirmed in the phase III trial (79%), despite the absence of dose limiting toxicity (DLT) conventionally defined by Grade 3 and 4 events. Dose escalation in the absence of tolerability concerns would be possible to implement in confirmatory studies and in clinical practice. However, implementation of lower starting doses (with the option to be increased as needed) is currently challenging due clinicians' perception that patients could unacceptably be exposed to subtherapeutic doses.

The potential to remove the need to dose to MTD was discussed during the workshop, particularly with the newer targeted therapies for which the dosing rationale should be different from what was implemented for nonselective chemotherapeutics. Further consideration should be given to alternative methods such as Bayesian methods, which could be coupled with PK/PD modeling to identify biologically effective dose from preclinical and emerging clinical data would allow the therapeutic index to be optimized in early clinical development.

The case of osimertinib, an epidermal growth factor receptor inhibitor, was used to illustrate how pr-clinical PK/PD relationships can be integrated with clinical PK variability using a model-based approach, to answer the question of "what dose should we aim to escalate to in the clinic to match efficacious exposure in preclinical models." A mathematical model relating PK, PD, and efficacy had been developed in animals for osimertinib during the discovery program.[93] This model was used to place differences between mouse and human PK into context (including a simulated clinical dose response in humans). The approach of integrating clinical PK with a preclinical PD-efficacy

relationship provided a way to augment early clinical data with a richer pr-clinical dataset and allowed the biologically effective dose (based on animal models and clinical exposure) to be identified. However, more research is still needed on the ongoing question of quantitative translation of animal models of cancers to the clinic.

Antibiotics

Bacterial load can generally not be quantitatively measured over time in patients and clinical concentration-bacterial killing data for antibiotics is therefore rarely available, with very few exceptions. The treatment period with antibiotic drugs is typically relatively short, but scheduling is often important to rapidly clear the bacteria and hence minimize the development of resistance. Because of the lack of clinical dose-ranging data, the exposure–response relationship for antibiotics usually relies on preclinical data. Good correlation to clinical response has been demonstrated for many classes of antibiotics. Antibiotics directly treat the cause of the disease, the bacteria, and thus studies in preclinical systems and model-based translation pose a unique opportunity in this therapeutic area. It also means that bacterial killing should be the same across different indications, as long as penetration into the tissue site is taken into account. The time-course of growth, drug-induced bacterial killing, and emergence of resistant bacteria is investigated using *in vitro* systems with static or dynamic antibiotic concentrations. Studies in animals are typically focused on a single timepoint for evaluation of drug effects and include mouse models of sepsis, pneumonia, and urinary tract infections, and the most commonly applied model: the thigh infection model.[94]

PD modeling of *in vitro* data can provide valuable information on the rate of killing and the risk and rate of takeover of resistance. This dynamic information, however, is rarely coupled with clinical PK in a mechanistic manner, even though the potential of mechanism-based PK/PD-models for selection of dosing regimens for antibiotics has been demonstrated in several published studies.[95,96] An aspect of this is the importance of and potential need to account for "host" and immune competence.

In current practice, a PK/PD target is instead defined based on the minimum inhibitory concentration of the bacteria (MIC), a summary measure of the PK profile (e.g., AUC, Cmax, and time above the MIC), and a fixed timepoint response in a mouse infection model. Stochastic simulations from a population PK model are performed to compute the proportion of patients that reaches the target (the probability of target attainment (PTA)), which can be compared between different dosing regimens. Dose selection can be based on a prediction of the majority of patients reaching PK/PD target attainment.

An example that illustrated the latter approach was presented during the workshop for phase III dose selection of fixed dose combination of ceftazidime-avibactam β-lactam–β-lactamase inhibitors in nosocomial pneumonia, complicated intraabdominal (cIAI), and complicated intrauterine (cUTI) infections. After target plasma and site of effect concentrations and PK/PD index were established using *in vitro* (hollow fiber) and animal data, PTA at the most critical times were derived from PK modeling of phase I and II studies and using Monte Carlo simulations to calculate PTA in the intended phase III populations. Phase III dose was selected based on a prediction of $>$90% PTA across indications. The data prospectively collected in phase III trial confirmed the adequacy of the dose selected.

Immunology drugs

As far as immunosuppressive drugs are concerned, the range of doses explored during drug development may be limited due to ethical considerations, given the well-documented PK variability and need for concentration controlled studies. In organ transplantation development programs, typical dose-finding studies are difficult, and often the dose rationale which has shown potentially adequate efficacy in a relatively small study is selected for assessment in a phase III trial. Under such conditions, it is challenging to ascertain whether the dosing regimen is optimal with respect to benefit/risk. Such assessments are even more complicated by the fact that, in this context, the drugs are not at PK steady state during a nonnegligible portion of the analysis period, and when drugs are administered in combinations, it is challenging to distinguish the relative contribution of individual drugs.[97] In such situations, model-based characterization of the exposure–response relationship may complement or even substitute the incomplete dose–response information.

A case study was presented during the workshop, illustrating the value of an exposure–response analysis in supporting the combination therapy of tacrolimus and everolimus in liver transplantation. Given that a placebo control arm was deemed unethical and in the absence of historical placebo data, it was not possible to estimate the contribution of everolimus to the overall efficacy of the combination therapy by means of traditional analyses (i.e., directly or indirectly via a noninferiority mechanism). Instead, an exposure–response approach was used to estimate the actual contribution of everolimus to the overall immunosuppressive response.

Dose finding in special populations

The importance and the challenges of an appropriate characterization of DER for children and the elderly were discussed during the workshop. This is particularly critical, knowing that the elderly are often the main users of drugs currently developed and the challenges to develop products in pediatrics, a complex population for which growth and organ maturation is key to consider. For both populations, age is rarely an independent source of variability, but correlates with other factors such as changes in physiology that directly impact PK/PD. Of particular note is that these populations may potentially be more at risk in terms of safety and efficacy—therefore the need for accurate dose recommendations is reinforced. The potential for changes in PK or PD should be assessed, such as a change in organ capacity (e.g., renal failure) or a comedication which could lead to drug–drug interaction risks; or loss of reserve capacity (e.g., CNS receptors), increased sensitivity (e.g., increased bleeding risk with anticoagulants), changes in cardiovascular function, or comorbidity.

Table 1 Main characteristics of advanced methods for phase II/III study design optimization and phase III dose selection as discussed during the EMA/FPIA workshop

Method	Recommended use during drug development	Assumptions	Requirements for use	Challenges and limitations
Empirical models (dose-responses)	Phase III Dose selection	-Time invariant dose-response -The optimal dose is within the range of the dose tested - Absence of unconsidered covariates for dose response	Wide enough range of doses to be explored	Ignorance of or very minimal assumptions of underlying pharmacology and determinants of response to treatment
Multiple comparison procedures and modeling (MCP-Mod)	Phase III Dose selection	-Time invariant dose-response -The optimal dose is within the range of the dose tested - Absence of unconsidered covariates for dose response - Phase II patients are representative of the target population	Adequate model prespecification	Does not apply when very different irregular treatment regimens are compared Does not apply to dose escalation studies -Unanticipated trends in the data
Model averaging	Phase III Dose selection	-Time invariant dose-response -The optimal dose is within the range of the dose tested - Absence of unconsidered covariates for dose response -Phase II patients are representative of the target population	Adequate model prespecification	Does not apply when very different irregular treatment regimens are compared Does not apply to dose escalation studies -Unanticipated trends in the data
Pharmacometrics model	Phase II/III study design optimization and Phase III Dose selection	-Phase II patients are representative of the target population	Relevant data for adequate characterization of dose-exposure-response are available	Based on specific model assumptions
Quantitative systems pharmacology models	Phase II/III study design optimization and Phase III Dose selection		Relevant pathways from dose to clinical response are well characterize on both time and space scales	Does not apply when the mechanisms of actions and determinants of drug response are unknown
Fisher Information Matrix-based methods	Phase II/III study design optimization	The prespecified model does not deviate too much from the actual model	Relevant models to be used as basis for optimization are available	
Clinical trial simulations	Phase II/III study design optimization	The prespecified model does not deviate too much from the actual model	Available model to be used as basis for simulation	Time and resource demanding
Adaptive dose ranging	Phase II/3 study design optimization		Prespecification of models to be used and decisions related to different scenarios	Time and resource demanding

Modelling and simulation (M&S) is a tool that provides the possibility for explicit generalized learning on the impacts of development and aging on disease and pharmacology. M&S can aid in scaling the DER from a general adult population to a special population, whether pediatrics or the elderly, and propose dose adjustments. Such a discussion is in line with the EMA's Geriatric Medicines Strategy[98] and the ongoing discussion in EMA on extrapolation.

GENERAL DISCUSSION AND IMPORTANT RECOMMENDATIONS FROM THE WORKSHOP

There are many methods available to support the design and analysis of dose-ranging studies. These methods are not mutually exclusive and should be used in conjunction. Main features, applicability, and limitations of different methods are summarized in **Table 1.** Multidisciplinary dialog is needed for optimal used of the different methods.

The limitations of the pairwise comparisons were acknowledged by the different stakeholders during the workshop. While model-based approaches were considered as relatively superior, the selection of the model to be used in a particular situation is not a straightforward and easy decision to make. Assumptions, context for use, and limitations of different methods need to be accounted for. The main considerations are summarized in **Table 1.**

Both empirical models developed using a data-driven approach and pharmacology-based models integrating the available knowledge on the drug and the disease from different sources (*in vitro*, animal, and previous human studies, etc.) were presented. Empirical cross-sectional dose–response models are valuable when the doses tested cover the relevant dose range and when there is no relevant time and/or covariate effects on the shape of the DER relationship. Another implicit assumption when using empirical dose–response models is that variability due to drug PK is less important than that due to the PD: dose is therefore considered as an acceptable exposure marker when describing the response. Lastly, the tested population is assumed representative of the target population in all the aspects relevant to the dose–response.

When pharmacology/physiology guided methods (PMx, QSP) are used, covariate effects can be assessed, which constitutes a clear added-value. When QSP methods are used, determinants of treatment outcomes are characterized in an even more robust manner, due to the direct link to the underlying physiological and pathological processes. When model-based approaches are used for data analysis, model assumptions should decrease with increasingly available data, as shown in **Figure 2**. In terms of the level of data needed for the analysis and confidence in the outcome, pairwise comparisons and systems pharmacology models constitute the two extremes cases, whereas empirical functions and pharmacometrics lie in between.

Model selection carries an inherent risk of choosing a model that is not an optimal approximation of the DER relationship. Integrative methods such as MCP-mod and model averaging aim to select the dose for phase III trials based on objective weighting of a set plausible candidate models and

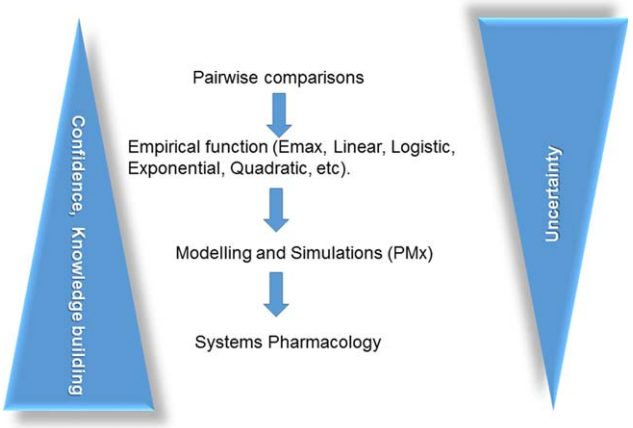

Figure 2 Schematic representation of comparison between methods with regard to assumption, uncertainty, and knowledge building.

related parameters, based on acceptable fitting of the available data. From a regulatory viewpoint, dose selection was clearly identified as a "shared risk" during the workshop and not solely a sponsor's risk. Design factors such as dose range, study durations, number of patients included, timing, and number of PK/PD/response measures and other inclusion/exclusion criteria should be carefully considered during drug development and appropriate methods such as FIM-based methods, CTS, and adaptive design used as needed. Importantly, dose selection should be performed with a benefit/risk assessment in mind. Optimal dosing in both target and subgroups of patients at high risk for lack of efficacy and/or for safety should be an inherent part of the dose-finding exercise. This is particularly true for drugs with high variability in their PK and PD. Dose adjustment should be considered for some subgroups of patients, as needed. The need for and the value of early dialog with regulators using the available procedures such as scientific advice, protocol assistance, qualification opinion, and advice were highlighted during the workshop.[99,100101]

Some areas for further discussions have been identified. They include the possibility for a solid phase II and a confirmatory phase III could be sufficient in a registration dossier, the need to ensure that we make the general adult development data informative for pediatric and geriatric developments, how we can support the generation of system data to better inform our understanding of sources to variability in DER, and the need for industry to have more regulatory support, in addition to the ICH E4 guideline, to facilitate improved study design and regulatory agreement on their clinical program.

Disclaimer. The content of this article reflects the authors' own views and should not be considered as recommendations from the European Medicines Agency, the Food and Drug Administration, or any European national medicines agency.

Acknowledgment. Martin Posch was supported by the European Union's Seventh Framework Programme for research, technological

development, and demonstration under grant agreement no 603160 (Asterix).

Conflict of Interest. The authors declare no conflicts of interest.

1. Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **3**, 711–715 (2004).
2. Burock, S., Meunier, F. & Lacombe, D. How can innovative forms of clinical research contribute to deliver affordable cancer care in an evolving health care environment? *Eur. J. Cancer.* **49**, 2777–2783 (2013).
3. Freidlin, B., Korn, E.L., Gray, R. & Martin, A. Multi-arm clinical trials of new agents: some design considerations. *Clin. Cancer Res.* **14**, 4368–4371 (2008).
4. Kaitin, K.I. & DiMasi, J.A. Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000–2009. *Clin. Pharmacol. Ther.* **89**, 183–188 (2011).
5. Onakpoya, I.J., Heneghan, C.J. & Aronson, J.K. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med.* **14**, 10 (2016).
6. Senn, S. *Statistical Issues in Drug Development*, 2nd ed. (Chichester, Wiley, 2007).
7. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002877.pdf>.
8. Holford, N.H. The target concentration approach to clinical drug development. *Clin. Pharmacokinet.* **29**, 287–291 (1995).
9. Gurwitz, J.H. Start low and go slow: Dosing of antipsychotic medications in elderly patients with dementia. *Arch. Intern. Med.* **155**, 2017–2018 (1995).
10. Ebling, W.F. & Levy, G. Population pharmacodynamics: strategies for concentration- and effect-controlled clinical trials. *Ann. Pharmacother.* **30**, 12–19 (1996).
11. Holford, N.H., Kimko, H.C., Monteleone, J.P. & Peck, C.C. Simulation of clinical trials. *Annu. Rev. Pharmacol. Toxicol.* **40**, 209–234 (2000).
12. Holford, N., Black, P., Couch, R., Kennedy, J. & Briant, R. Theophylline target concentration in severe airways obstruction — 10 or 20 mg/L? A randomised concentration-controlled trial. *Clin. Pharmacokinet.* **25**, 495–505 (1993).
13. Kraiczi, H., Jang, T., Ludden, T. & Peck, C.C. Randomized concentration-controlled trials: motivations, use, and limitations. *Clin. Pharmacol. Ther.* **74**, 203–214 (2003).
14. Lledo-Garcia, R., Hennig, S. & Karlsson, M.O. Comparison of dose-finding designs for narrow-therapeutic-index drugs: concentration-controlled vs. dose-controlled trials. *Clin. Pharmacol. Ther.* **86**, 62–69 (2009).
15. Holford, N.H.G. & Buclin, T.M.D. Safe and effective variability — a criterion for dose individualization. *Ther. Drug Monit.* **34**, 565–568 (2012).
16. Held, J. *et al.* Ferroquine and artesunate in African adults and children with Plasmodium falciparum malaria: a Phase 2, multicentre, randomised, double-blind, dose-ranging, non-inferiority study. *Lancet Infect. Dis.* **15**, 1409–1419 (2015).
17. Pocock, S.J. Clinical Trials—A Practical Approach (Chichester, Wiley, 1983).
18. Thomas, N., Sweeney, K. & Somayaji, V. Meta-analysis of clinical dose response in a large drug development portfolio. *Stat. Biopharm. Res.* **6**, 302–317 (2014).
19. Thomas, N. & Dooti, R. Analysis of clinical dose–response in small-molecule drug development: 2009–2014. *Stat. Biopharm. Res.* doi 10.1080/19466315.2016.1256229 (2016) [Epub ahead of print].
20. Kallen, A. *Computational Pharmacokinetics* (CRC Press, Boca Raton, FL; 2007).
21. Bretz, F., Pinheiro, J.C. & Branson, M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics* **61**, 738–748 (2005).
22. Pinheiro, J., Bornkamp, B., Glimm, E. & Bretz, F. Model-based dose finding under model uncertainty using general parametric models. *Stat. Med.* **33**, 1646–1661 (2014).
23. CHMP (2014) 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. <http://www.ema.europa.eu/docs/en_GB/document_library/-Regulatory_and_procedural_guideline/2014/02/WC500161027.pdf>.
24. FDA (2016) Determination Letter: <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM508700.pdf>.
25. Bornkamp, B. *et al.* Innovative approaches for designing and analyzing adaptive dose-ranging trials. *J. Biopharm. Stat.* **17**, 965–995 (2007).
26. Mercier, F., Bornkamp, B., Ohlssen, D. Wallstroem, E. Characterization of dose–response for count data using a generalized MCP-Mod approach in an adaptive dose-ranging trial. *Pharm. Stat.* **14**, 359–367 (2015).
27. Verrier, D., Sivapregassam, S. & Solente, A.C. Dose-finding studies, MCP-Mod, model selection, and model averaging: two applications in the real world. *Clin. Trials.* **11**, 476–484 (2014).
28. Buckland, S.T., Burnham K.P. & Augustin N.H. Model selection: an integral part of inference. *Biometrics* **53**, 603–618 (1997).
29. <http://www.page-meeting.org/pdf>_assets/272-AOKI_HOOKER_modelAveragePoster.pdf>.
30. <http://www.page-meeting.org/?abstract=3121>.
31. Draper, D. Discussion of "Decision models in screening for breast cancer" by G. Parmigiani. In *Bayesian Statistics 6* (eds. J.M. Bernardo, J. Berger, P. Dawid, & A.F.M. Smitheds) 541–543 (Oxford University Press, Oxford, UK; 1999).
32. Raftery and Zheng <http://www.stat.washington.edu/www/research/reports/2003/tr433.pdf> (2003).
33. Breiman, L. Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996).
34. Schorning, K., Bornkamp, B., Bretz, F. & Dette, H. Model selection versus model averaging in dose finding studies. *Stat. Med.* **35**, 4021–4040 (2016).
35. Hoeting, J.A., Madigan, D., Raftery, A.E. & Volinsky, C.T. Bayesian model averaging: a tutorial (with comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors). *Statist. Sci.* **14**, 382–417 (1999).
36. Grasela, T.H. & Slusser, R. Improving productivity with model-based drug development: an enterprise perspective. *Clin. Pharmacol. Ther.* **88**, 263–268 (2010).
37. Kimko, H. & Pinheiro, J. Model-based clinical drug development in the past, present and future: a commentary. *Br. J. Clin. Pharmacol.* **79**, 108–116 (2015).
38. Lalonde, R.L. *et al.* Model-based drug development. *Clin. Pharmacol. Ther.* **82**, 21–32 (2007).
39. Milligan, P.A. *et al.* Model-based drug development: a rational approach to efficiently accelerate drug development. *Clin. Pharmacol. Ther.* **93**, 502–514 (2013).
40. ICH: Dose-response information to support drug registration: 1994. <http://webcache.googleusercontent.com/search?q=cache:CFVmkSHtt6gJ:www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E4/Step4/E4_Guideline.pdf>+&cd=1&hl=en&ct=clnk&gl=uk. Accessed 15 June 2016.
41. Vozeh, S. *et al.* The use of population pharmacokinetics in drug development. *Clin. Pharmacokinet.* **30**, 81–93 (1996).
42. Aarons, L. *et al.* Population approaches in drug development, report of an expert meeting to discuss population pharmacokinetic/pharmacodynamic software. *Eur. J. Clin. Pharmacol.* **46**, 389–391 (1994).
43. Wu, K. *et al.* Population pharmacodynamic model of the longitudinal FEV1 response to an inhaled long-acting anti-muscarinic in COPD patients. *J. Pharmacokinet. Pharmacodyn.* **38**, 105–119 (2011).
44. Bender, B.C., Schindler, E. & Friberg, L.E. Population pharmacokinetic-pharmacodynamic modeling in oncology: a tool for predicting clinical response. *Br. J. Clin. Pharmacol.* **79**, 56–71 (2015).
45. Quartino, A.L., Friberg, L.E. & Karlsson, M.O. A simultaneous analysis of the time-course of leukocytes and neutrophils following docetaxel administration using a semi-mechanistic myelosuppression model. *Invest. New Drugs* **30**, 833–845 (2012).
46. Miller, R. *et al.* Exposure-response analysis of pregabalin add-on treatment of patients with refractory partial seizures. *Clin. Pharmacol. Ther.* **73**, 491–505 (2003).
47. Lunn, D.J., Wakefield, J. & Racine-Poon, A. Cumulative logit models for ordinal data: a case study involving allergic rhinitis severity scores. *Stat. Med.* **20**, 2261–2285 (2001).
48. Wang, Y. *et al.* Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. *Clin. Pharmacol. Ther.* **86**, 167–174 (2009).
49. Fedorov, V., Wu, Y. & Zhang, R. Optimal dose-finding designs with correlated continuous and discrete responses. *Stat. Med.* **31**, 217–234 (2012).
50. Bretz, F., Dette, H. & Pinheiro, J.C. Practical considerations for optimal designs in clinical dose finding studies. *Stat. Med.* **29**, 731–742 (2010).
51. Dayneka, N.L., Garg, V. & Jusko, WJ. Comparison of four basic models of indirect pharmacodynamic responses. *J. Pharmacokinet. Biopharm.* **21**, 457–478 (1993).
52. Mager, D.E., Wyska, E. & Jusko, WJ. Diversity of mechanism-based pharmacodynamic models. *Drug Metab. Dispos.* **31**, 510–518 (2003).
53. Aarons, L. *et al.* Role of modeling and simulation in Phase I drug development. *Eur. J. Pharm. Sci.* **13**, 115–122 (2001).
54. Kaneko, M. *et al.* Confirmation of model-based dose selection for a Japanese phase III study of rivaroxaban in non-valvular atrial fibrillation patients. *Drug Metab. Pharmacokinet.* **28**, 321–331 (2013).
55. Zhang, X. *et al.* Dose selection using a semi-mechanistic integrated glucose-insulin-glucagon model: designing Phase 2 trials for a novel oral glucokinase activator. *J. Pharmacokinet. Pharmacodyn.* **40**, 53–65 (2013).
56. Marshall, S., Macintyre, F., James, I., Krams, M. & Jonsson, N.E. Role of mechanistically-based pharmacokinetic/pharmacodynamic models in drug development: a case study of *a therapeutic protein. Clin. Pharmacokinet.* **45**, 177–197 (2006).
57. Ren, Y.P. *et al.* Model-based meta-analysis of the effects of non-selective and α1-selective GABAA receptor agonists in healthy volunteers. *Eur. J. Clin. Pharmacol.* **71**, 1209–1221 (2015).
58. Soger, *et al.* Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms. An NIH White Paper by the QSP Workshop Group - October, 2011.
59. Barrett, J.S., Fossler, M.J., Cadieu, K.D. & Gastonguay, M.R. Pharmacometrics: a multidisciplinary field to facilitate critical thinking in drug development and translational research settings. *J. Clin. Pharmacol.* **48**, 632–649 (2008).
60. Aldridge, B.B. *et al.*, Physicochemical modeling of cell signalling pathways. *Nat. Cell Biol.* **8**, 1195–1203 (2006).
61. Mesens, N. *et al.* Screening for phospholipidosis induced by central nervous drugs: comparing the predictivity of an in vitro assay to high throughput in silico assays. *Toxicol. In Vitro* **24**, 1417–1425 (2010).

62. Rodriguez, B. *et al.* The systems biology approach to drug development: application to toxicity assessment of cardiac drugs. *Clin. Pharmacol. Ther.* **88**, 130–134 (2010).
63. Atkinson, A.C. & Donev, A.N. Optimum Experimental Designs (Oxford University Press, Oxford, UK; 1992).
64. Retout, S. & Mentré, F. Optimization of individual and population designs using Splus. *J. Pharmacokinet. Pharmacodyn.* **30**, 417–443 (2003).
65. Ogungbenro, K. & Aarons, L. Population Fisher information matrix and optimal design of discrete data responses in population pharmacodynamic experiments. *J. Pharmacokinet. Pharmacodyn.* **38**, 449–469 (2011).
66. Dette, H., Pepelyshev, A. & Wong, WK. Optimal designs for dose-finding experiments in toxicity studies. *Bernoulli* **15**, 124–145 (2009).
67. Fedorov, V.V. & Wu, Y. Dose finding designs for continuous responses and binary utility. *J. Biopharm. Stat.* **17**, 1085–1096 (2007).
68. Fedorov, V.V. & Leonov, S.L. Optimal design of dose response experiments: a model-oriented approach. *Drug Inf. J.* **35**, 1373–1383 (2001).
69. Deleon, A.C.P. & Atkinson, A.C. Optimum experimental-design for discriminating between 2 rival models in the presence of prior information. *Biometrika* **78**, 601–608 (1991).
70. Atkinson, A.C. & Bogacka, B. Compound D- and D-s-optimum designs for determining the order of a chemical reaction. *Technometrics* **39**, 347–356 (1997).
71. Ogungbenro, K., Graham, G., Gueorguieva, I. & Aarons, L. The use of a modified Fedorov exchange algorithm to optimise sampling times for population pharmacokinetic experiments. *Comput. Methods Programs Biomed.* **80**, 115–125 (2005).
72. Amstein, R. *et al.* RIDO:Multimedia CD-Rom software for training in drug development via PK/PD principles and simulation of clinical trials. *Pharm. Res.* **13**, S452 (1996).
73. Bhattaram, V.A. *et al.* Impact of pharmacometrics on drug approval and labeling decisions: a survey of 42 new drug applications. *AAPS J.* **7**, E503–512 (2005).
74. Bland, J.M. Computer simulation of clinical trial as an aid to teaching the concept of statistical significance. *Stat. Med.* **5**, 193–197 (1986).
75. Boessen, R., Knol, M.J., Groenwold, R.H. & Roes, K.C. Validation and predictive performance assessment of clinical trial simulation models. *Clin. Pharmacol. Ther.* **89**, 487–488 (2011).
76. Bonate, P.L. Clinical trial simulation in drug development. *Pharm. Res.* **17**, 252–256 (2000).
77. Chan, P.L., Nutt, J.G. & Holford, N.H. Levodopa slows progression of Parkinson's disease: external validation by clinical trial simulation. *Pharm. Res.* **24**, 791–802 (2007).
78. Gastonguay, M., Pentikis, H., Alexander, M. & Lee, L. Applying modeling and simulation to improve the design of clinical efficacy trial. *Clin. Pharmacol. Ther.* **65**, 181 (1999).
79. Girard, P., Cucherat, M. & Guez, D. Clinical trial simulation in drug development. *Therapie* **59**, 297–304 (2004).
80. Gobburu, J.V. & Marroum, P.J. Utilisation of pharmacokinetic-pharmacodynamic modeling and simulation in regulatory decision-making. *Clin. Pharmacokinet.* **40**, 883–892 (2001).
81. Hale, M., Gillespie, W.R., Gupta, S.K., Tuk, B. & Holford, N.H.G. Clinical trial simulation — streamlining your drug development process. *Appl. Clin. Trials* **5**, 35–40 (1996).
82. Holford, N., Ma, S.C. & Ploeger, B.A. Clinical trial simulation: a review. *Clin. Pharmacol. Ther.* **88**, 166–182 (2010).
83. Holford, N.H., Kimko, H.C., Monteleone, J.P. & Peck, C.C. Simulation of clinical trials. *Annu. Rev. Pharmacol. Toxicol.* **40**, 209–234 (2000).
84. Jumbe, N., Yao, B., Rovetti, R., Rossi, G. & Heatherington, A.C. Clinical trial simulation of a 200-microg fixed dose of darbepoetin alfa in chemotherapy-induced anemia. *Oncology (Huntingt)* **16**(10 suppl. 11), 37–44 (2002).
85. Kimko, H.C. & Duffull, S.B. *Simulation for Designing Clinical Trials. A Pharmacokinetic-Pharmacodynamic Modeling Perspective* (ed. Swarbrick, J.) (Marcel Dekker, New York, 2003).
86. Lockwood, P., Ewy, W., Hermann, D. & Holford, N. Application of clinical trial simulation to compare proof-of-concept study designs for drugs with a slow onset of effect; an example in Alzheimer's disease. *Pharm. Res.* **23**, 2050–2059 (2006).
87. EMA. Report from Dose Finding Workshop. <http://www.ema.europa.eu/docs/en_GB/document_library/Report/2015/04/WC500185864.pdf> (2015).
88. FDA Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics.
89. <http://www.ema.europa.eu/docs/en_GB/document_library/Other/2011/02/WC500102291.pdf>.
90. Elsäβer, A. *et al.* Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. *Trials* **15**, 383 (2014).
91. Posch, M. *et al.* Testing and estimation in flexible group sequential designs with adaptive treatment selection. *Stat. Med.* **24**, 3697–3714 (2005).
92. Dragalin, V.V.F.Y.W. Two-stage design for dose-finding that accounts for both efficacy and safety. *Stat. Med.* **27**, 5156–5176 (2008).
93. Yates, J.W.T. *et al*, Irreversible inhibition of EGFR: Modeling the combined pharmacokinetic-pharmacodynamic relationship of osimertinib and its active metabolite AZ5104. *Mol. Cancer Ther.* **15**, 2378–2387 (2016).
94. Ambrose, P.G. *et al.* Pharmacokinetics-Pharmacodynamics of antimicrobial therapy: It's not just for mice anymore. *Clin. Infect. Dis.* **44**, 79–86 (2007).
95. Nielsen, E.I. & Friberg, L.E. Pharmacokinetic-pharmacodynamic modeling of antibacterial drugs. *Pharmacol. Rev.* **65**, 1053–1090 (2013).
96. Laouénan, C., Guedj, J. & Mentré, F. Clinical trial simulation to evaluate power to compare the antiviral effectiveness of two hepatitis C protease inhibitors using nonlinear mixed effect models: a viral kinetic approach. *BMC Med. Res. Methodol.* **13**, 60 (2013).
97. Dumortier, T. *et al.* Estimating the contribution of everolimus to immunosuppressive efficacy when combined with tacrolimus in liver transplantation: a model-based approach. *Clin. Pharmacol. Ther* **97**, 411–418 (2015).
98. <http://www.ema.europa.eu/docs/en_GB/document_library/Other/2011/02/WC500102291 .pdf>.
99. <http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000049.jsp&mid=WC0b01ac05800229b9>.
100. <http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000319.jsp&mid=WC0b01ac0580022bb0>.