TUTORIAL

Power Determination During Drug Development: Is Optimizing the Sample Size Based on Exposure-Response Analyses Underutilized?

Aksana K. Jones[1,†], Ahmed Hamed Salem[1,2] and Kevin J. Freise[1,*]

The use of model-based drug development (MBDD) has been demonstrated to improve the efficiency of clinical trial design. However, MBDD complexity can limit its use, particularly early in clinical development. In this tutorial, a simple and generalizable exposure-response analysis approach to determine the power for dose-ranging studies is presented and described. We identified situations in which higher power and sample size reduction is achieved by utilizing the exposure-response powering methodology compared with conventional power calculations.

CLINICAL TRIAL PLANNING

Drug development struggles with high costs and time-consuming processes.[1] Budget and time restrictions can limit the number of subjects to be enrolled and/or number of doses to be examined. These restrictions can lead to limited information to make informed decisions for further development.

Interestingly, the use of model-based drug development (MBDD) analysis has been demonstrated to have the potential of drastically reducing the required study size in phase II clinical trials and thereby reducing costs, time spent, and the number of patients exposed to an investigational drug.[1] Furthermore, utilizing dose-ranging studies during phase I trials offers important design advantages in the form of dose-response information to inform dose selection in later studies. Previous work has also shown that using prior knowledge from published data to evaluate the power of studies is valuable in the planning of clinical trials and evaluation of study protocols.[2] Rekić et al.[3] demonstrated that utilizing MBDD achieves more precise dose–response characterization and, hence, facilitates decision making and dose selection for phase II trials. The importance of applying exposure-response analyses throughout clinical drug development as tools for understanding efficacy and safety of investigational therapies and as quantitative support for decision making has been outlined in the tutorial by Overgaard et al.[4]

Overall, it is shown that during the planning of a clinical dose-ranging study, adequate powering and choosing the number and ranges of doses to be studied are critical parts of the study design and the subsequent determination of dosing reccomend ations for future studies. In this tutorial, we outline an exposure-response based approach to determine power for dose-ranging studies. The exposure-response powering methodology is compared with conventional power calculations and situations in which clear reductions in sample size can be achieved are identified.

POWER AND SAMPLE SIZE

A power analysis can determine the minimum sample size required to reject the null hypothesis with an assumed effect size between two dose groups with a certain probability. Multiple factors affect the power of a study, including the variability in response, type I error-rate (i.e., α), statistical test, and effect size. Therefore, the assumptions used for determining power during the study planning process are essential.

Conventional methodology

Assume a study is being planned to evaluate the efficacy of a drug at two doses with the response being a binary end point (responses: yes or no). For conventional power calculations, one would assume the following hypothesis:

$$H_0 : P_1 = P_2 \text{ vs. } H_a : P_1 \neq P_2 \tag{1}$$

where H_0 is the null hypothesis, P_1 is the probability of response in dose group 1, P_2 is the probability of response in dose group 2, and H_a is the alternative hypothesis.

Given a type I error-rate α, the number of subjects in each treatment group (n) and the probabilities of the responses in the treatment groups P_1 and P_2, respectively, the power can be calculated as:

$$1 - \beta \left(\alpha, P_1 - P_2 \right) = 1 - \left(\frac{Z_{1-\frac{\alpha}{2}} \frac{\sqrt{2\bar{P}(1-\bar{P})}}{\sqrt{n}} - (P_1 - P_2)}{\frac{\sqrt{P_1(1-P_1)+P_2(1-P_2)}}{\sqrt{n}}} \right) \tag{2}$$

where $\bar{P}$ is the average of the expected rates in the treatment groups, $\bar{P} = \frac{P_1 + P_2}{2}$. The critical value of $Z_{1-\frac{\alpha}{2}}$ is taken from the normal distribution table.

Exposure-response methodology

Let us still assume that the probability (P) of response for two dose groups is P_1 and P_2. If the clinical pharmacokinetics

[1]Abbvie, Inc., North Chicago, Illinois, USA; *Correspondence: Kevin J. Freise (kevin.freise@abbvie.com)
[2]Faculty of Pharmacy, Ain Shams University, Egypt
[†]Present address: Cognigen, Buffalo, NY.

(PK) of the drug is also known from the first-in-human study, then the corresponding drug exposures (e.g., area under the concentration-time curve (AUC)) can be defined for the respective dose groups 1 and 2.

Then based on the logit transformation, defined as:

$$\text{logit}(P) \overset{\text{def}}{=} \log\left(\frac{P}{1-P}\right) \tag{3}$$

the intercept (β_0) and slope (β_1) of the logistic regression equation are given by:

$$\beta_0 = \text{logit}(P_1) - \beta_1 \cdot \text{AUC}_1, \tag{4a}$$

$$\beta_1 = \frac{\text{logit}(P_2) - \text{logit}(P_1)}{\text{AUC}_2 - \text{AUC}_1} \tag{4b}$$

with

$$P(\text{AUC}) = 1/\left(1 + e^{-(\beta_0 + \beta_1 \cdot \text{AUC})}\right) \tag{5}$$

where AUC_1 and AUC_2 are the typical AUC values for dose groups 1 and 2, respectively. Any exposure metric is appropriate for this methodology, but as part of the motivating example we will only consider AUC. Use of the typical exposure metric, instead of the full distribution, for calculating β_0 and β_1 in Eqs. 4a and 4b may result in (small) differences in the predicted probability due to the nonlinear relationship in Eq. 5, although the difference is generally expected to be minimal. The exposure-response framework, H_0 and H_a, can be defined as:

$$H_0:\beta_1 = 0 \text{ vs. } H_a:\beta_1 \neq 0 \tag{6}$$

Thus, H_0 is rejected if the slope term of the exposure-response model is significantly different from 0. Although the motivating example assumes a binary end point with a logistic regression exposure-response relationship, it can be seen from Eq. 6 that any response end point or exposure-response relationship could be considered as long as the single parameter is definable that isolates the drug effect (i.e., β_1).
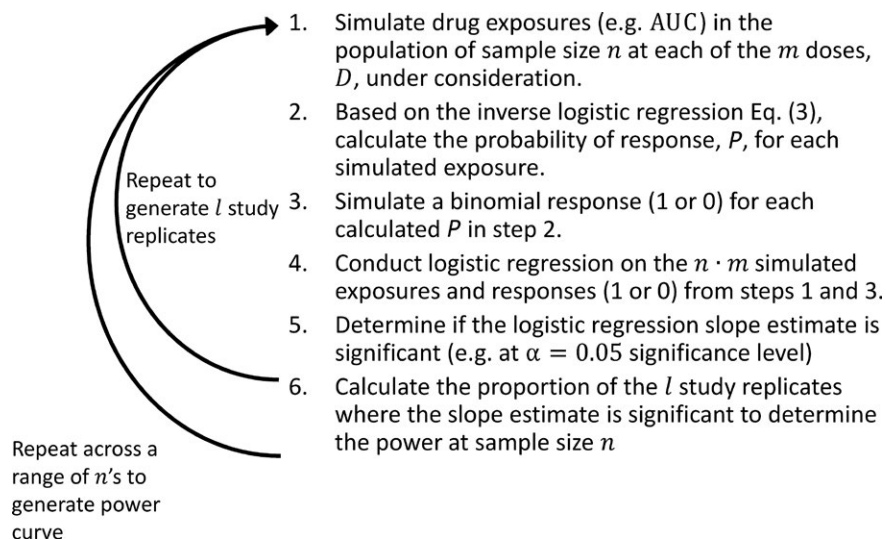
In order to determine the power using exposure-response methodology, the distribution of the drug exposure (i.e., PK) in the population at a given dose, D, is also needed. This information will need to be available from phase I studies evaluating the clinical PK of the drug. Hence, let us assume based on the previous developed population PK model, AUC exposures can be obtained based on the following distribution:

$$\text{AUC} = \frac{D}{\text{CL/F}}, \quad \text{CL/F} \sim \log N(\log(\theta), \omega^2) \tag{7}$$

where CL/F is the apparent clearance of the drug, which is assumed to be log-normally distributed with population mean θ and variance ω^2 on the logarithmic scale.

Given the relationships outlined in Eqs. 5 and 7 and the hypotheses defined in Eq. 6, the power at sample size n at each of m doses, D, under consideration can be determined by simulation using the algorithm outlined in **Figure 1**. The probability of response, P, for each simulated exposure can be obtained based on the exposure-response relationship in our example logistic regression (Step 2, **Figure 1**). For each calculated probability, the response can be simulated (Step 3, **Figure 1**), and an exposure-response analysis on the $n \cdot m$ simulated exposures and responses can be conducted (Step 4, **Figure 1**). Then, at a given significance level, determine if the exposure-response relationship is significant (Step 5, **Figure 1**). Lastly, the proportion of the l study replicates where the exposure-response relationship is significant is to be calculated to determine the power at sample size n (Step 6, **Figure 1**). This can be repeated to generate/study replicates across a range of sample size n's to generate power curves and, thus, determine the sample size needed to show whether there is a significant difference across each of the m doses, D.

In this tutorial, the type I error-rate (i.e., α, level of significance) was assumed at 5%, and the statistical power, the probability of identifying a real difference (i.e., $1 - \beta$, type II error-rate), was aimed to be 80%.

1. Simulate drug exposures (e.g. AUC) in the population of sample size n at each of the m doses, D, under consideration.

2. Based on the inverse logistic regression Eq. (3), calculate the probability of response, P, for each simulated exposure.

Repeat to generate l study replicates

3. Simulate a binomial response (1 or 0) for each calculated P in step 2.

4. Conduct logistic regression on the $n \cdot m$ simulated exposures and responses (1 or 0) from steps 1 and 3.

5. Determine if the logistic regression slope estimate is significant (e.g. at $\alpha = 0.05$ significance level)

6. Calculate the proportion of the l study replicates where the slope estimate is significant to determine the power at sample size n

Repeat across a range of n's to generate power curve

Figure 1 General workflow for conducting sample size simulations. AUC, area under the concentration-time curve.

Table 1 Summary of simulation results

Scenario	Treatment	N per group	N for study	Simulated power
A	Conventional power	65	130	80.5
	Exposure-response	45	90	82.2
B Different slopes	1 and 2 mg – β_1 = 0.5 mL/µg	150	300	77.4
	1 and 2 mg – β_1 = 1 mL/µg – Reference	50	100	87.8
	1 and 2 mg – β_1 = 2 mL/µg	30	60	86.8
C Different intercepts	1 and 2 mg – β_0 = –3	55	110	80.4
	1 and 2 mg – $\beta 0$ = –1.5 – Reference	45	90	81.0
	1 and 2 mg – β_0 = –0.5	60	120	80.2
D Number of doses	One dose: 2 mg	155	155	82.3
	Two doses: 1 and 2 mg – Reference	45	90	81.6
	Three doses: 1, 2, and 3 mg	20	60	90.7
E Dose range	1 and 2 mg – Reference	45	90	82.2
	1.5 and 3 mg	30	60	83.3
	0.5 and 3.5 mg	15	30	83.3
F Varying CVs	1 and 2 mg – 10% CV	60	120	82.2
	1 and 2 mg – 25% CV	45	90	82.2
	1 and 2 mg – 40% CV	30	60	80.7
	1 and 2 mg – Conventional Power Calc.	65	130	80.5
Ixazomib case study scenario 1	3 and 4 mg	40	80	86.2
	3 and 4 mg – Conventional Power Calc.	770	1540	80.1
Ixazomib case study scenario 2	Two doses: 3 and 4 mg	40	80	86.2
	Three doses: 3, 4, and 5 mg	20	60	84.1
Ixazomib case study scenario 3	3 and 4 mg: 42.3% CV	40	80	86.2
	3 and 4 mg: 10% CV	230	460	84.4

CV, coefficient of variation.

METHOD EVALUATION

The algorithm outlined in **Figure 1** was followed to generate power curves with *n* subjects ranging from 10–150 subjects, *m* dose groups ranging from 2–3, and *l* = 1,000 study replicates. The R-script used to generate the power curves is included in **Supplementary Material S1**.

Factors influencing logistic regression exposure-response power

Multiple scenarios were simulated to evaluate the factors that impact the power of exposure-response analyses using a logistic regression model. The reference scenario for the simulation represents intercept β_0 of –1.5, slope β_1 of 1 mL/µg, apparent drug clearance CL/F of 1 L/hour, a coefficient of variation (CV) of 25%, and doses of 1 and 2 mg. The following scenarios were simulated:

- Comparing the power curve obtained from the conventional methodology and the exposure-response methodology
- Comparing power curves for different slopes (0.5, 1 (reference), and 2 mL/µg) to evaluate the impact of the steepness of the exposure-response relationship
- Comparing power curves for different intercepts (–3, –1.5 (reference), and –0.5) to evaluate the impact of the observed response in the absence of drug (e.g., placebo effect)

- Comparing power curves for different number of doses (1 (2 mg, reference), 2 (1 and 2 mg, reference), and 3 (1, 2, and 3 mg))
- Comparing power curves for different dose ranges (1.5 and 3 mg, 1 and 2 mg (reference), and 0.5 and 3.5 mg).
- Comparing power curves for different CVs (10%, 25%, and 40%) to evaluate the impact of PK variability

Case study: ixazomib dose ranging

Following completion of the ixazomib dose-ranging proof-of-concept study, ixazomib doses of both 3 and 4 mg once weekly were considered to likely provide favorable risk/benefit ratios with estimated clinical benefit rates of 37% and 43%, respectively.[5] A hypothetical subsequent phase II study was considered, which seeks to determine if the clinical benefit rate at a 4 mg dose is superior to a 3 mg dose. Combining the above information with the known population PK of ixazomib, the calculations to determine the intercept and slope logistic regression exposure-response model parameters are as follows:

At a dose of 3 mg (D_1 = 3 mg), the assumed response rate is 37% (i.e., P_1 = 0.37) and at a dose of 4 mg (D_2 = 4 mg), the assumed response rate is 44% (i.e., P_2 = 0.44).[5] Based on the established population PK model,[6] the ixazomib typical CL/F is 2.0 L/hour with 42.3% CV, and the corresponding AUCs at 3 and 4 mg doses are 0.09 and 012 µg·h/
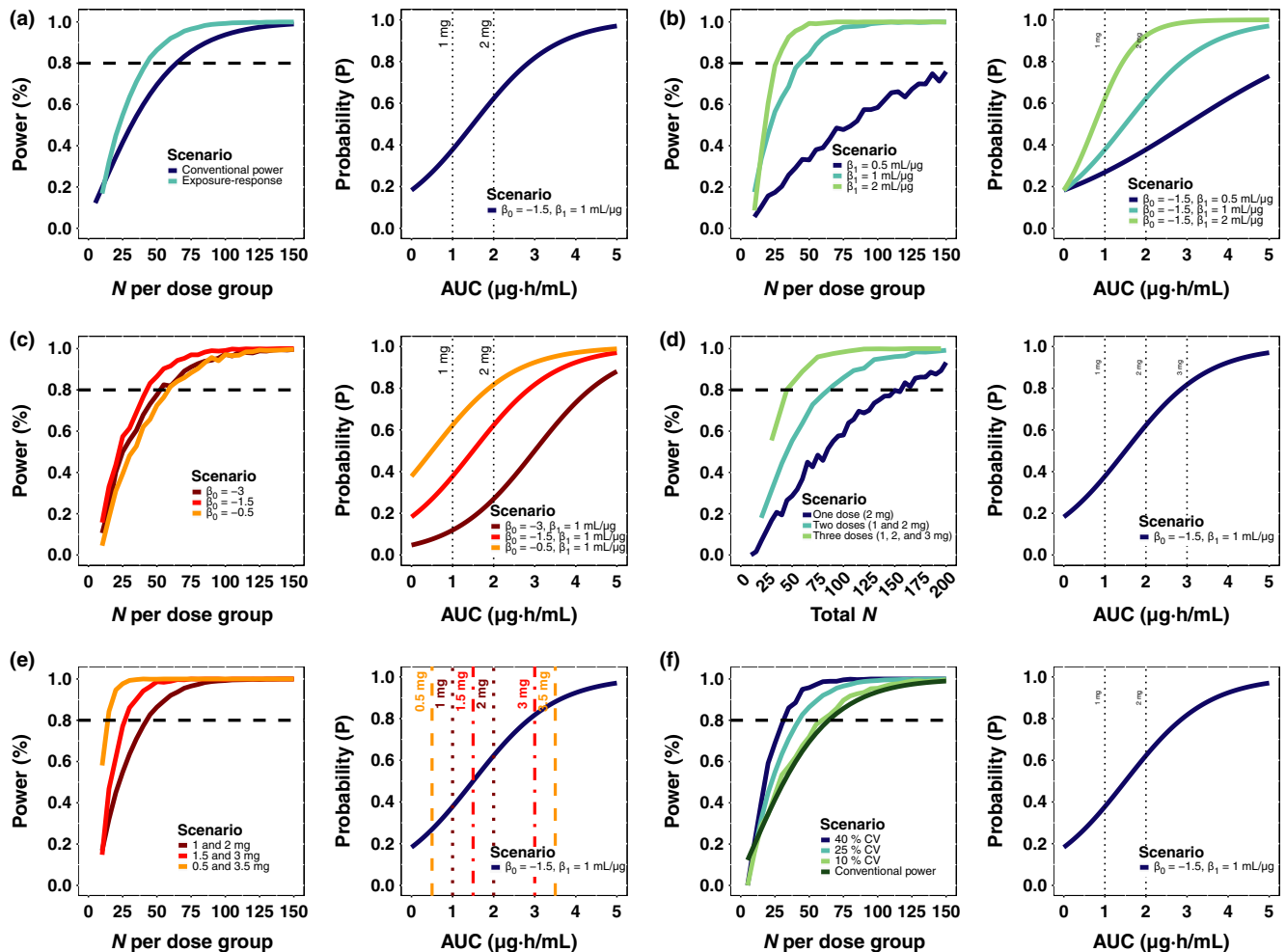
Figure 2 Simulation of power curves of influencing factors. (**a**) Conventional and exposure-response methodology, (**b**) exposure-response slope (0.5, 1 (reference), and 2 mL/μg), (**c**) exposure-response intercept (−3, −1.5 (reference), and −0.5), (**d**) number of doses (1 (2 mg), 2 (1 and 2 mg, reference), and 3 (1, 2, and 3 mg)), (**e**) dose range (1.5 and 3 mg, 1 and 2 mg (reference), and 0.5 and 3.5 mg), and (**f**) pharmacokinetic coefficient of variation (CV; 10, 25 (reference), and 40%). Left-hand and right-hand plots in each panel display the power curves and exposure-response relationships, respectively. The reference scenario for the simulation represents intercept β_0 of −1.5, slope β_1 of 1 mL/μg, apparent drug clearance CL/F of 1 L/hour, a CV of 25%, and doses of 1 and 2 mg. AUC, area under the concentration-time curve.

mL, respectively (i.e., AUC_1 and AUC_2, respectively). Using Eqs. 4a and 4b, estimates of β_0 and β_1 using (P_1, AUC_1) and (P_2, AUC_2) can be calculated ($\beta_0 = -2.03$, $\beta_1 = 15.12$). From these, the algorithm in **Figure 1** can be used with $n = 10$ to $n = 150$ subjects, $m = 2$ dose groups, and $l = 1,000$ study replicates to determine the exposure-response power curve.

Subsequently, the power curves were compared using the exposure-response powering methodology vs. the conventional powering methodology, and study design features to improve the exposure-response power were explored.

Three scenarios were simulated to represent the case study, specifically:

- Base scenario comparing two dose groups (3 vs. 4 mg) with the conventional power calculation
- Comparing two dose groups (3 and 4 mg) and three dose groups (3 vs. 4 vs. 5 mg)

- Comparing power curves for different CVs (10% and 42.3%)

RESULTS

Table 1 summarizes the number of subjects per study required to achieve at least 80% power for all the scenarios tested in this tutorial.

Factors influencing logistic regression exposure-response power

The different factors influencing the logistic regression exposure-response powering are displayed in **Figure 2**. It is shown when comparing the exposure-response powering methodology with the conventional powering methodology that at a given sample's size, the exposure-response

methodology has a higher power (**Figure 2a**). Additionally, both the slope (β_1) and intercept (β_0) terms of the logistic regression model influenced the power (**Figure 2b,c**, respectively). The relationship between slope and power was monotonic, with the greatest power occurring with the highest slope term. Contrary, the power varied with the intercept term in a nonmonotonic way. Collectively, it can be observed that the power depends on the exposures being in a region with the greatest change in the probability of response at a given change in exposure.

Increasing the number of doses and range of the doses examined improved the exposure-response power at a given sample size (**Figure 2d,e**, respectively). Both design features result in spacing the observations across a greater range of exposures, improving the precision to estimate the slope parameter. The impact of spacing the observations across a greater exposure range is also demonstrated by the exposure-response power simulations that examined the impact of clearance CV (**Figure 2f**), with the higher clearance CV resulting in a greater range of exposures. It is also evident that as the CV clearance decreases the advantage of the exposure-response powering methodology over conventional powering methodology disappears, with little, if any advantage at the lowest CV of 10% in this simulation.

Case study: ixazomib dose ranging

Figure 3a shows the resulting power curve and demonstrates the improved power using exposure-response powering methodology compared with conventional methodology for our case study. In fact, using conventional methodology determining if 4 mg ixazomib once weekly is superior to 3 mg once weekly is likely prohibitive for a phase II trial, with > 770 patients per dose group needed to achieve 80% power, despite a 7% improvement in response rate. In contrast, the sample size needed to achieve 80% power using exposure-response powering methodology is about $n = 40$ patients per dose group (80 patients total). To further improve the power using exposure-response powering methodology, a third dose of 5 mg ixazomib, which is still less than the maximum tolerated dose of 5.5 mg,[5] was added to the study design. **Figure 3b** depicts power curves using exposure-response powering methodology for the three dose groups compared with two dose groups. With the addition of the third dose group at 5 mg ixazomib, the total sample size (n per group) is reduced from 80 (40) subjects to 60 (20) subjects. It should be noted that similar results can be achieved with the addition of a lower third dose of 2 mg ixazomib. Further reductions in total sample size are possible, if additional doses across a larger range are also added to the study design. Assuming a CV of 10% for ixazomib shows the increase in sample size from 40 subjects per dose group to 215 subjects per dose group (**Figure 3c**), showing also with this case example that as the CV clearance decreases, the sample size required to achieve 80% power increases.

DISCUSSION

A simple and generalizable exposure-response analysis approach to determine the power for dose-ranging studies was presented that integrates seamlessly with the MBDD paradigm. The operating characteristics of one of the most common types of clinically applied exposure-response models, a logistic regression model, were explored. With this type exposure-response it was demonstrated that an exposure-response powering methodology can be used to help guide the planning of clinical trials and substantially increase the power of clinical trials and/or decrease the number of subjects required within a trial. Although such concepts have been presented before within the context of MBDD,[1,3,4] the use may have been limited by previously perceived complexity in its use. The primary advantage of MBDD analysis for guiding the planning of clinical trials is the reduction in sample size when utilizing exposure-response powering methodology.

The evaluation of the factors influencing the exposure-response powering methodology revealed the following key attributes: Overall, it was shown that the higher the variability in exposure, the less subjects are required for showing a significant difference between the dose groups (in our example, 30 vs. 120 subjects to be enrolled into the study for variabilities of 40% CV and 10% CV, respectively). This is in alignment with the ability to define an exposure-response relationship better if the range of exposures is wider. Furthermore, if the exposure-response relationship is more pronounced, the sample size will decrease as the difference between the two dose groups will be clearer. Lastly, if more dose groups are tested (i.e., from two dose groups to three dose groups), resulting in a wider range of exposures on the exposure-response curve, less subjects are required. Additionally, the increase in dose groups will also lead to a more informative re-estimation of the exposure-response relationship given the wider range of exposures achieved with a higher range of dose groups. This is true if the drug exposures are not saturated and the exposures achieved are not at the maximum of the exposure-response relationship. Hence, it is important to understand the exposure-response relationship as early as possible. Further increases in power can be achieved if modeling longitudinal end points are considered.[1–3] However, the increased model complexity involved in longitudinal end points limits the use of such methodology due to the increased time it takes to implement and greater difficulty in being able to communicate exactly what (e.g., model parachute) you are powering the study on.

These powering methods can also be directly applied to dose-response analyses. Some of the factors that indicate when dose-response analyses may be more useful than exposure-response analyses for dose selection are reviewed by Hsu (2009) and Berges and Chen (2013).[7,8] Specifically, the exposure-response analyses were substantially more precise than dose-response methods when clearance was highly variable between individuals, but the former suffered when the exposure measurement error was high. Consistent with the previous investigations on dose selection precision, the exposure-response powering methodology had the greatest advantage over conventional power methodology when the intersubject variability in exposure was higher (**Figure 2f**). Similar to both dose-response and exposure-response dose determination,[7] an important performance driver for the exposure-response powering methodology was the dose range evaluated (**Figure 2d,e**).
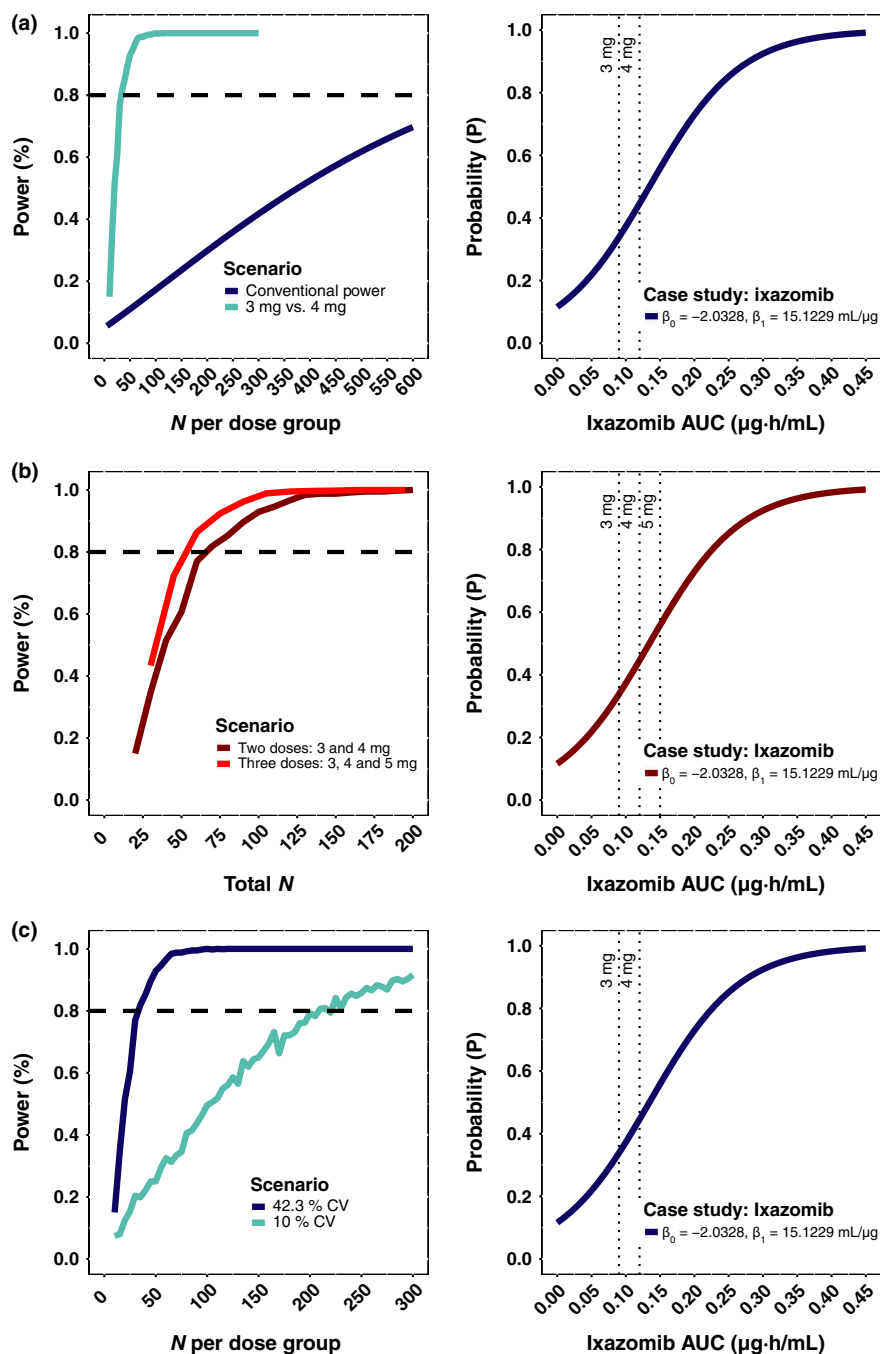
Figure 3 Simulation of power curves. Case study ixazomib (adapted from Gupta *et al.*[5]). (**a**) Conventional and exposure-response methodology, (**b**) number of doses (2 (3 and 4 mg, reference) and 3 (3, 4, and 5 mg)), (**c**) pharmacokinetic coefficient of variation (CV; 42.3 (reference) and 10%). Left-hand and right-hand plots in each panel display the power curves and exposure-response relationships, respectively. AUC, area under the concentration-time curve.

The case study described in this tutorial did not include a control group. This is more frequently seen in development of agents for cancers with unmet medical need. However, the presented approach can be easily generalized to placebo or active, controlled dose-ranging studies. One limitation of the proposed exposure-response power methodology is that drug exposure is reduced to a single value. To account for factors such as partial noncompliance with dosing regimen, sensitivity analyses may need to be completed where the exposure metric, such as AUC, is an average value reflecting observed dose intensity rather than a steady-state value. In other cases, such as comparing once daily with twice daily dosing, exposure metrics, such as AUC, would need to be avoided as it would not reflect the different dosage regimens. A maximum or trough concentration would likely be a more appropriate exposure

metric in this instance. Another limitation of the proposed exposure-response power methodology is that it assumes the true exposure-response relationship follows the modeled exposure-response relationship. Due to model misspecification, the exposure-response relationship used for powering and modeling will not be completely accurate, likely resulting in some overestimation of the trial power. Simulation-based sensitivity analyses using a variety of "true" exposure-response relationships, while fitting using the assumed (logistic regression) exposure-response model, can be used to assess the fragile ness of the power determination to model misspecification.

Overall, exposure-response modeling has demonstrated its utility in dose selection and trial design during drug development and is being used more and more throughout clinical trial planning for early stages in drug development.[9] We have demonstrated that utilizing exposure-response powering methodology has the potential of reducing the required study size in dose-ranging clinical trials and thereby reducing costs, time spent, and the number of patients exposed to an investigational drug. The exposure-response relationship utilized in the MBDD analysis can be easily borrowed from drugs in the same class or from other development programs within the same therapeutic area. It should be noted that other than typically readily available information on the clinical PK of the drug, minimal additional assumptions have been needed to define the necessary information for exposure-response powering compared with conventional powering methodology.

Use of exposure-response powering methodology compared with conventional powering methodology is similar to designing first-in-patient proof-of-concept (POC) human trials using concentrated vs. distributed inference space.[10] In concentrated designs, only placebo and the highest tolerated dose are evaluated in the POC study in order to make a (Go/No-Go) decision on whether to continue to advance an experimental therapy. With only two treatment groups and only one test group, conventional powering methodology would be used. In contrast, distributed designs evaluate multiple doses across a range in the POC study in order to both make a (Go/No-Go) decision on whether to advance an experimental therapy and to gain information on the dose-response relationship. As such, the exposure-response powering methodology utilizes a distributed design that not only can improve study power compared with conventional powering methodology but also provides useful information on the exposure (dose)-response relationship. It should be noted that simply demonstrating that an exposure-response relationship is statistically different from 0 for powering purposes is most appropriate for phase IIa or POC studies, in which the objective is to determine if a drug is effective or not. For phase IIb studies, in which the objective is to determine the most appropriate dose to take into a phase III study, more precision on the exposure (dose)-response relationship would be required than simply powering for an effect.

One of the frequent limitations to exposure-response based analyses is when subjects' exposures are observed following administration of various doses (i.e., without randomization of subjects to predefined exposures).[11,12] Observing (nonrandomized) exposures of course is the most common way of obtaining exposure-response information and the one proposed to be used within for the described exposure-response power analysis. However, as exposure is (in part) an outcome in this situation, biases in the estimated exposure-response relationship due to confounding factors may occur. In these instances, additional diagnostics and sensitivity analyses can be used to support the analyses and identify any bias that may have an impact on interpretations.[11]

MBDD analysis could be limited though by not having a clear defined exposure-response relationship. Lacking the exposure-response relationship of an investigational drug will lead back to utilize conventional power calculations for the planning of the dose-ranging study. This shows again the importance of determining the exposure-response relationship as early as possible in clinical development to guide not only dosing decisions but also guide the planning of clinical trials in regard to sample size.

SUMMARY

This tutorial has shown the primary advantage of utilizing MBDD analysis, which is to reduce sample size when utilizing exposure-response powering methodology for guiding the planning of clinical trials. By providing a relatively simple, and readily generalizable, exposure-response analysis approach to determine power for dose-ranging studies, the exposure-response powering methodology was compared with conventional power calculations and was shown to enable clear reductions in sample size. The detailed factors influencing the power and relative simplicity in the outlined approach will hopefully overcome resistance and increase uptake of methodology that stands to substantially improve the efficiency of drug development. We believe that this tutorial is helpful during the planning of dose-ranging clinical trials and potentially led to reduction of the required study size and thereby reducing costs, time spent, and the number of patients exposed to an investigational drug during drug development. Furthermore, we hope that this tutorial not only outlines the approach for utilizing exposure-response powering methodology but also gives insights on when exposure-response methodology increases the information to make informed decisions.

1. Karlsson, K.E., Vong, C., Bergstrand, M., Jonsson, E.N. & Karlsson, M.O. Comparisons of analysis methods for proof-of-concept trials. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e23 (2013).

2. Bhattaram, V.A., Siddiqui, O., Kapcala, L.P. & Gobburu, J.V.S. Endpoints and analyses to discern disease-modifying drug effects in early Parkinson's disease. *AAPS J.* **11**, 456–464 (2009).

3. Rekić, D., Röshammar, D. & Simonsson, U.S.H. Model based design and analysis of phase II HIV-1 trials. *J. Pharmacokinet. Pharmacodyn.* **40**, 487–496 (2013).

4. Overgaard, R.V., Ingwersen, S.H. & Tornøe, C.W. Establishing good practices for exposure-response analysis of clinical endpoints in drug development. *CPT Pharmacometrics Syst. Pharmacol.* **4**, 565–575 (2015).

5. Gupta, N., Labotka, R., Liu, G., Hui, A.M. & Venkatakrishnan, K. Exposure-safety-efficacy analysis of single-agent ixazomib, an oral proteasome inhibitor, in relapsed/refractory multiple myeloma: dose selection for a phase 3 maintenance study. *Invest. New Drugs* **34**, 338–346 (2016).

6. Gupta, N., Zhao, Y., Hui, A.M., Esseltine, D.L. & Venkatakrishnan, K. Switching from body surface area-based to fixed dosing for the investigational proteasome inhibitor ixazomib: a population pharmacokinetic analysis. *Br. J. Clin. Pharmacol.* **79**, 789–800 (2015).

7. Hsu, C.H. Evaluating potential benefits of dose-exposure-response modeling for dose finding. *Pharm. Stat.* **8**, 203–215 (2009).

8. Berges, A.C.C. Dose finding by concentration-response versus dose-response: a simulation-based comparison. *Eur. J. Clin. Pharmacol.* **69**, 1391–1399 (2013).

9. Gibbs, J.P. Prediction of exposure-response relationships to support first-in-human study design. *AAPS J.* **12**, 750–758 (2010).

10. Dodds, M.G., Salinger, D.H., Mandema, J., Gibbs, J.P. & Gibbs, M.A. Clinical trial simulation to inform phase 2: comparison of concentrated vs. distributed first-in-patient study designs in psoriasis. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e58 (2013).

11. Nedelman, J.R., Rubin, D.B. & Sheiner, L.B. Diagnostics for confounding in PK/PD models for oxcarbazepine. *Stat. Med.* **26**, 290–308 (2007).

12. US Food and Drug Administration (FDA). Exposure-Response Relationships – Study Design, Data Analysis, and Regulatory Applications. <https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm072109.pdf>. Published 2003. Accessed November 9, 2018.