ORIGINAL ARTICLE

Clinical Trial Simulation to Inform Phase 2: Comparison of Concentrated vs. Distributed First-in-Patient Study Designs in Psoriasis

MG Dodds[1], DH Salinger[1], J Mandema[2], JP Gibbs[1] and MA Gibbs[1]

Clinical trial simulation (CTS) and model-based meta-analysis (MBMA) can increase our understanding of small, first-in-patient (FIP) trial design performance to inform Phase 2 decision making. In this work, we compared dose-ranging designs vs. designs testing only placebo and the maximum dose for early decision making in psoriasis. Based on MBMA of monoclonal antibodies in the psoriasis space, a threshold of greater than a 50 percentage point improvement over placebo effect at the highest feasible drug dose was required for the advancement in psoriasis. Studies testing only placebo and the maximum dose made the correct advancement decision marginally more often than dose-ranging designs in the majority of the cases. However, dose-ranging studies in FIP trials offer important design advantages in the form of dose–response (D–R) information to inform Phase 2 dose selection. CTS can increase the efficiency and quality of drug development decision making by studying the limitations and benefits of study designs prospectively.

Research and development within the pharmaceutical industry has become increasingly challenging, as evidenced by late-stage development failures due to lack of efficacy and problems securing regulatory approval for new drugs.[1] Model-based drug development is a successful alternative to the existing empirical drug development path.[2–4] This mathematical approach to analyzing pharmacokinetic and pharmacodynamic data has proved particularly useful in the efficient selection of dosages in patient studies.[5] More recently, it has been recognized that applying modeling efforts earlier in the drug development pathway can lead to more efficient clinical trials by (i) including rich, high-quality data from Phase 1 studies in the analysis of subsequent trials, (ii) including biomarker and efficacy endpoints early in the development process to design subsequent trials, and (iii) performing clinical trial simulations (CTSs) to predict patient outcome before subsequent trials are conducted.[6] CTS requires information on the expected outcomes of a clinical trial, and model-based meta-analysis (MBMA) can provide that by pooling data across studies to define the mean and variance associated with a measured characteristic across studies. In a drug development setting, MBMA techniques have been used to characterize the dose–response (D–R) relationship for statins, biologics for use in inflammatory bowel disease and rheumatoid arthritis, antimigraine treatments, trabectedin, and latanoprost.[7–10] In addition, MBMA has been utilized to define relationships between biomarkers and efficacy endpoints using quantitative models.[11,12] Quantitative models support rational dose selection and inform trial design decisions with the objective of maximizing the probability of success, as recently demonstrated in the treatment of diabetes with rivoglitazone.[13] Additional information derived from disease and trial models are elements of model-based drug development.[2] More recently, the drug industry has focused on quantitative pharmacology at early stages to inform Phase 2 dose selection.[14]

Safety, tolerability, and pharmacokinetics are often the primary objectives of first-in-human studies. While these studies may show that high doses (e.g., a 700 mg intravenous infusion of a monoclonal antibody) are safe and well tolerated, study outcomes may lack biomarker or clinical signs due to the healthy volunteer population studied. Understanding the portfolio value of novel compounds is a primary objective of a first-in-patient (FIP) study, where biomarker or clinical signs may be available to guide drug development with a relatively small investment of time and resources. Results from this early study may provide insight into whether a sponsor should accelerate or slow development and/or resource allocation based on an early look at the exposure–response relationship. Evidence from such a study can be useful in making decisions around progression to and design of the next stage of clinical development (large-scale, patient dose-ranging studies). Sponsors may categorize FIP studies as Phase 1 or early Phase 2; rather than using a particular numerical term, we refer to this study by its function: the first study of a drug in a relevant patient population. Two key pieces of information are desired from FIP studies. First, the design should demonstrate unambiguous efficacy related to the disease process so that a sound decision on proceeding, or not, into further clinical development, the so-called "Go/No-Go decision" (G/NG), can be made. Second, the design should reveal the relationship between drug dose level and efficacy that shapes choices of dose and regimen in further clinical development, the so-called "dose–response relationship".

However, when first-in-human (healthy volunteer) studies demonstrate no changes in biomarker or clinical signs, it is not clear what study design should be employed for FIP

[1]Department of Pharmacokinetics & Drug Metabolism, Amgen, Seattle, Washington, USA; [2]Quantitative Solutions, Menlo Park, California, USA. Correspondence: MA Gibbs (gibbsm@amgen.com)

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds et al.

2

studies. For the purposes of this work, it is assumed that safety and tolerability of a drug up to 700 mg via the intravenous route have been established in healthy volunteers prior to considering a FIP study. In addition, no dose-selection guidance from biomarker or clinical signs are available in healthy volunteers. That is, the drug under consideration only has meaningful and detectable biologic impact in a disease state, and no safety signals at very high doses of drug are evident. If, instead, some measure of biologic impact via biomarker, clinical signs, or safety signals were available, this D–R information could be used to further tailor the FIP study design under consideration.[15] One typical FIP study design may be to assign the patient either placebo or the highest tolerated dose in healthy volunteers (hereafter: concentrated designs). Alternatively, patients may be assigned different dose levels across the range of acceptable dose levels tolerated in healthy volunteers (hereafter: distributed design). Further, it is assumed that a small ($N = 16$) FIP trial in psoriasis patients will be used to demonstrate clinical impact. In the context of psoriasis trials, the key metric around which decisions are taken is the improvement in Psoriasis Area Severity Index (PASI) in patients during therapy.[16] Early G/NG and D–R can be derived from single doses of biologics by monitoring PASI over a 12-week window after a single injection.[17] Traditionally, a sponsor may elect to move into a larger, dose-ranging (Phase 2) study in patients. It is the purpose of this work to determine if it is possible to use a smaller ($N = 16$) FIP study to gain insight into the feasibility of a drug before moving into larger and more costly studies. Finally, it is assumed that positive results for this FIP study are required for further drug development in Phase 2.

MBMA can inform trial design by leveraging available data to provide a realistic basis for CTS. Assuming a lack of biomarker or clinical signs in first-in-human for a candidate drug, we use CTS to understand if a FIP trial design can correctly identify the properties of successful drugs ("good drugs") and unsuccessful drugs. In general, a FIP trial design, at a minimum, should be able to correctly identify a successfully marketed drug as such with a low false-negative rate (i.e., failing to identify a successful molecule as such). In addition, MBMA can assist in creating other test cases ("bad drugs") where a

FIP trial design should detect poor performance with a low false-positive rate (i.e., failing to identify a poor molecule as such). To date, the application of meta-analysis to inform FIP study design has not been published.

The objective of this analysis was to examine the quality of G/NG and adequacy of estimated D–R given two FIP trial design strategies (concentrated vs. distributed designs). Efficient use of clinical trial subjects, time, and resources motivated this effort as an investigation of "learn and confirm" using trial simulation.[18] MBMA was utilized to provide realistic parameters for placebo response, drug potency, and maximal effect for a class of efficacious monoclonal antibodies with linear pharmacokinetics used in psoriasis which enabled subsequent trial simulation with the goal of optimizing FIP study design.

RESULTS
Test compound selection
MBMA of biologic treatments for psoriasis. Data from trials evaluating adalimumab, golimumab, and ustekinumab in patients with psoriasis were analyzed using MBMA.[19–23] The log of the mean ratio in PASI score over baseline was analyzed using a D–R meta-analysis, yielding model parameters for mean % change in PASI score (**Table 1**).[24] For the included marketed compounds, adalimumab, golimumab, and ustekinumab, the meta database analysis demonstrated that an E_{max} model was adequate to describe the D–R properties of these compounds. All of the marketed compounds were found to have equal maximum change in PASI % improvement (absolute maximal difference from placebo = 82.3%). The placebo rate was small but non-negligible. The marketed compounds were found to have different doses that produce half of the maximum effect (ED_{50} = 16.9, 45.5, and 13.9 mg for adalimumab, golimumab, and ustekinumab, respectively). **Figure 1** shows a plot of the expected PASI % improvement by dose for these marketed compounds and **Table 1** shows their E_{max} model parameters.

Additional test cases. Additional test cases were selected where E_0, ED_{50}, and E_{max} were perturbed from the ranges suggested by the marketed compounds to understand the

Table 1 Test case potency parameters and desired trial outcome

| Group | Compound | Simulation parameters | | Desired trial outcome | |
		Maximal absolute difference from placebo PASI % change (%)	ED_{50} (mg)	Correct G/NG	Estimated ED_{50} within
Marketed examples	adalimumab	82.3	16.9	G	8.50–33.8
	golimumab	82.3	45.5	G	22.8–91.0
	ustekinumab	82.3	13.9	G	6.90–27.7
No-Go examples	discontinumab	22.6	11.8	NG	5.90–23.6
	mehmimab	40.7	32.1	NG	16.0–64.1
Go examples	cuspmimab	58.8	32.1	G	16.0–64.1
	lowpomab	82.3	182	G	91.0–364
	nopomab	82.3	728	G	364–1460

Placebo response (E_0) is 9.5% for all compounds.
ED_{50}, the dose providing half maximal drug response; G/NG, Go/No-Go; PASI, Psoriasis Area Severity Index.

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
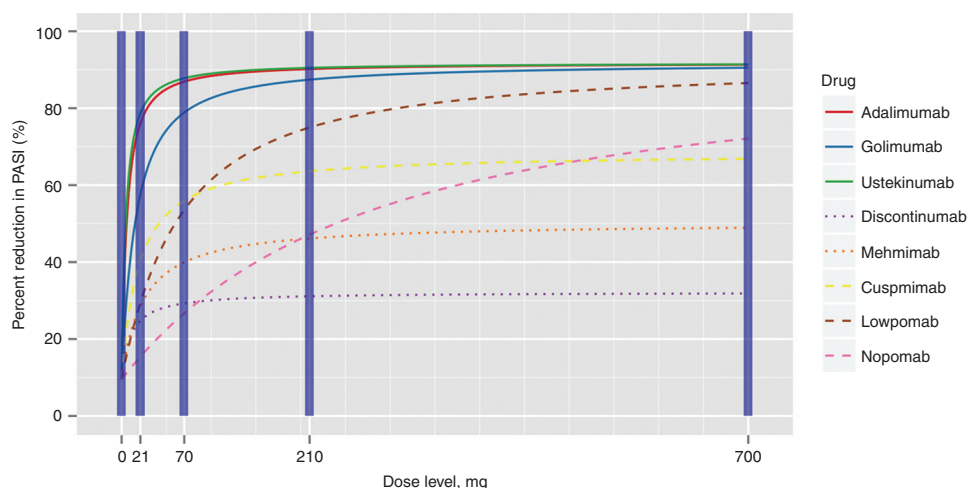Dodds *et al.*

npg

3

Figure 1 Dose–response relationships from a meta-analysis of marketed biologics adalimumab, golimumab, ustekinumab (solid lines); the hypothetical No-Go test cases discontinumab, mehmimab (short dashed lines); and the Go test cases cuspmimab, lowpomab, nopomab (long dashed lines). Expected typical geometric mean percent reduction in PASI (y-axis) vs. dose level (x-axis) are shown with the dose levels selected (0, 21, 70, 210, and 700 mg) for simulation study (blue shaded area) shown for reference.

design performance sensitivity to test compound properties (**Table 1**). Hypothetical cases, discontinumab and mehmimab, were selected with ED_{50} within the range of the marketed test cases, 11.8 and 32.1 mg, respectively, but with poor and marginal maximal absolute difference from placebo values of 22.6 and 40.7% assigned to discontinumab and mehmimab, respectively. The lower maximal response is a realistic assumption in the development of therapeutics targeted to novel mechanisms. Cuspmimab was selected to have the same potency as mehmimab, but with slightly higher maximal absolute difference from placebo value of 58.8%. Lowpomab and nopomab were assigned the maximal absolute difference from placebo values equal to those of the marketed test cases, 82.3%, but assigned ED_{50} values that were higher than the marketed cases, 182 and 728 mg, respectively, to test sensitivity of the analysis to these potentially difficult to develop compounds. **Figure 1** shows the D–R relationship for these additional test cases (discontinumab, mehmimab, cuspmimab, lowpomab, and nopomab) and **Table 1** shows their E_{max} model parameters.

Design performance metrics

G/NG criteria. Based on the MBMA results of the marketed compounds, the G/NG criteria were defined as a 50% point improvement from placebo response at the maximum feasible dose. That is, considering the strength of marketed molecules in this space and a maximal feasible dose of 700 mg, a novel agent would need to exhibit at least a 50% point absolute improvement in PASI score over placebo to be considered for further development. Marketed molecules in the space achieve 81–82% absolute improvement in PASI over placebo response at the highest feasible dose (700 mg), and so these criteria set a relatively permissive advancement threshold. **Table 1** and **Supplementary Table S1** give the true values of maximal response, ED_{50} and placebo response used in CTS, and the G/NG decision for a FIP study to be deemed successful in providing good G/NG decision making.

A FIP study design should suggest continued development for the marketed test cases, adalimumab, golimumab, and ustekinumab, as these compounds demonstrate an 81.8, 81.0, and 81.9% absolute difference between drug effect at 700 mg and placebo effect, respectively. A clinical study that suggests discontinuing development of these efficacious molecules is a false-negative result. A FIP study should suggest discontinuing development for the discontinumab and mehmimab cases, having only 22.3 and 39.4% absolute difference between drug effect at 700 mg and placebo effect, respectively. A clinical study that suggests continuing development of these molecules is a false-positive result. A FIP study should suggest continuing development for the cuspmimab, lowpomab, and nopomab cases, having 57.3, 77.0, and 62.6% absolute difference between drug effect at 700 mg and placebo effect, respectively. As these effects are close to the G/NG criteria, we would expect false-negative results to be more common than observed for the marketed test compounds.

D–R criteria. We define good D–R information as deriving an ED_{50} estimate from the FIP trial within twofold of the true value. That is, if a small FIP study returns an estimate of ED_{50} within twofold of the truth, doses selected in the next stage of development (Phase 2 dose-ranging) are likely to efficiently deliver a refined estimate of ED_{50}. In other applications, the criteria of twofold could be relaxed or made more stringent, depending on the weight of this information for development decisions. **Table 1** gives the true values of ED_{50} used in CTS and the range within which an estimate of ED_{50} must fall for a FIP study to be deemed successful in providing good D–R information.

FIP trial design choice

Distributed trial designs. The distributed trial design involves assigning psoriasis patients across a range of feasible and practical drug dose levels. Here, we assume that the drug product is available in a 70 mg/ml presentation and choose

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds *et al.*

4

doses of 21 mg (0.3 ml), 70 mg (1 ml), 210 mg (3 ml), and 700 mg (10 ml). Given the ED_{50} estimates from the MBMA of marketed test cases (16.9, 13.9, and 45.5 mg), the 21 and 70 mg dose levels are likely to produce submaximal response, whereas the 210 and 700 mg doses are likely to product maximal response in psoriasis patients. The results of this design are more complex to analyze than for the concentrated design, requiring a model of the D–R relationship. The design questions involve both the ratio of placebo to actively treated patients as well as the number of patients assigned to each dose level. **Table 2** describes the patient allocation choices considered in this work for a small ($N = 16$) FIP trial. The distributed designs spread all experimental information on determining placebo and D–R across the range of allowed doses, and differ only in the number of patients assigned placebo or active doses (7:1, 3:1, 5:3 and 1:1 active:placebo).

Concentrated trial designs. The concentrated trial design involves assigning all psoriasis patients either placebo or the highest feasible drug dose level tested in healthy volunteers (700 mg). The results of this design are relatively easy to analyze by standard ANOVA techniques, and the only design question involves the ratio of placebo-treated patients to active drug-treated patients. **Table 2** describes the patient allocation choices considered in this work for a small ($N = 16$) trial. These designs concentrate all experimental information on determining patient response given placebo or the highest feasible dose, and the designs differ only in the number of patients assigned placebo or the highest feasible dose (7:1, 3:1, 5:3 and 1:1 active:placebo).

Impact of design choice on G/NG decision making

Marketed test cases. **Figure 2** shows the design performance for G/NG decision making for the marketed test cases adalimumab, golimumab, and ustekinumab, in which

the correct decision after FIP is to continue development. All designs frequently (93–100%) identified the marketed biologic test case's Go decision correctly (**Table 3**). The percentage of correct Go decisions was similar between designs with matched active:placebo ratios, differing by 1% (**Table 3**). Design performance was sensitive to placebo patient numbers for all designs, favoring designs with an equal balance in active and placebo-treated subjects (**Table 3**).

No-Go hypothetical test cases. **Figure 2** shows the design performance for G/NG decision making for the hypothetical test cases, discontinumab and mehmimab, in which the correct decision is to discontinue development due to lack of competitiveness. All designs frequently (64–95%) identified the correct No-Go decision for these hypothetical test cases (**Table 3**). An advantage to the concentrated design is evident in the percentage of correct No-Go decisions for these hypothetical test cases (**Table 3**), and the advantage was 3–6%. Design performance was sensitive to placebo patient numbers for all designs, favoring designs with an equal balance in active and placebo-treated subjects (**Table 3**).

Go hypothetical test cases. **Figure 2** shows the design performance for G/NG decision making for the hypothetical test cases cuspmimab, lowpomab, nopomab, cases in which the correct decision was to continue development. All designs frequently (58–98%) identified the correct Go decision for these hypothetical test cases (**Table 3**). For these hypothetical test cases, the relative advantage between distributed and concentrated design was more nuanced that the previous cases. For cuspmimab, distributed designs were better by 2–3% for all matched active:placebo ratios. **Figure 2** shows that the median of the estimates of absolute improvement over placebo deviated from the true value for the concentrated designs more so than for the distributed designs.

Table 2 Distributed and concentrated design dose level assignments (mg)

Subject number	Distributed designs				Concentrated designs			
	7 active: 1 placebo	3 active: 1 placebo	5 active: 3 placebo	1 active: 1 placebo	7 active: 1 placebo	3 active: 1 placebo	5 active: 3 placebo	1 active: 1 placebo
1	0	0	0	0	0	0	0	0
2	21	21	0	0	0	0	0	0
3	21	21	21	21	700	0	0	0
4	21	21	21	21	700	0	0	0
5	0	0	0	0	700	700	0	0
6	70	70	0	0	700	700	0	0
7	70	70	70	70	700	700	700	0
8	70	70	70	70	700	700	700	0
9	210	0	0	0	700	700	700	700
10	210	210	210	0	700	700	700	700
11	210	210	210	210	700	700	700	700
12	210	210	210	210	700	700	700	700
13	700	0	0	0	700	700	700	700
14	700	700	700	0	700	700	700	700
15	700	700	700	700	700	700	700	700
16	700	700	700	700	700	700	700	700

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
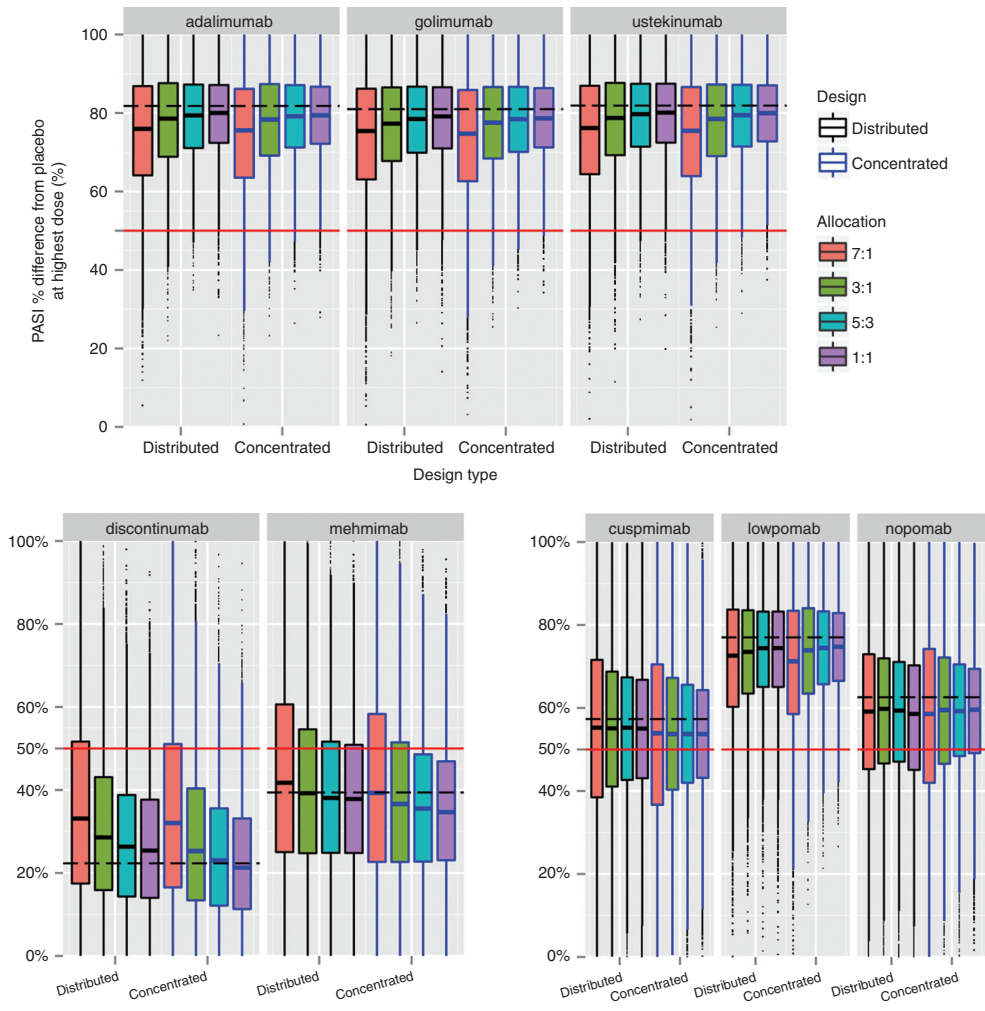Dodds *et al.*

npg

5

Figure 2 Design performance for Go/No-Go decision making for the marketed test cases adalimumab, golimumab, ustekinumab (top panel); the No-Go test cases discontinumab, mehmimab (lower left panel); and the Go test cases cuspmimab, lowpomab, nopomab (lower right panel). Estimates of PASI improvement over placebo at the highest dose are shown on the y-axis for each compound by design type (distributed and concentrated from left to right) and active to placebo ratios (from left to right) as boxplots. Go/No-Go criterion (PASI improvement over placebo at the highest dose > 50%) is shown as a solid, red, horizontal line and the true for simulation is shown as a dashed, black, horizontal line.

Table 3 Percentage of simulated trials providing the correct Go or No-Go development decision

Group	Drug	7 active:1 placebo		3 active:1 placebo		5 active:3 placebo		1 active:1 placebo	
		Distributed (%)	Concentrated (%)	Distributed (%)	Concentrated (%)	Distributed (%)	Concentrated (%)	Distributed (%)	Concentrated (%)
Marketed Examples	adalimumab	94	94	98	98	99	99	99	100
	golimumab	94	93	97	98	99	99	99	99
	ustekinumab	95	94	98	99	99	99	99	100
No-Go examples	discontinumab	78	81	87	90	91	94	92	95
	mehmimab	64	67	70	75	74	78	75	81
Go examples	cuspmimab	60	58	60	58	61	58	61	59
	lowpomab	90	88	94	95	95	97	94	98
	nopomab	68	66	69	69	70	72	66	73

For lowpomab and nopomab cases, distributed designs at the 7:1 active:placebo ratio were better by 2%, at the 3:1 active ratio equal to the concentrated designs, and at the 5:3 and 1:1 active:placebo ratios worse by 2–7%.

Impact of design choice on D–R information

Concentrated designs, by definition, offered no information regarding D–R. The results for the distributed designs are reported below.

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds *et al.*

6

Table 4 Percentage of simulated trials providing an ED_{50} estimate within two-fold of the true value

Group	Drug	7 active: 1 placebo (%)	3 active: 1 placebo (%)	5 active: 3 placebo (%)	1 active: 1 placebo (%)
Marketed examples	adalimumab	59	57	52	49
	golimumab	65	65	61	58
	ustekinumab	55	53	47	45
No-Go examples	discontinumab	18	17	14	14
	mehmimab	32	32	29	28
Go examples	cuspmimab	45	44	40	38
	lowpomab	52	53	51%	48
	nopomab	25	24	25	23

ED_{50}, the dose providing half maximal drug response.

Marketed test cases. All distributed designs frequently (45–65%) identified the marketed biologic test cases D–R correctly (**Table 4**). The success rate was ranked golimumab (ED_{50} = 45.5 mg) > adalimumab (ED_{50} = 16.9 mg) > ustekinumab (ED_{50} = 13.9 mg), which corresponds to the ED_{50} rank for these compounds (**Table 1**). Gaining an accurate ED_{50} estimate for these potent molecules was challenging, given the sample size (N = 16) and doses (21, 70, 210, and 700 mg) used in this simulation study. Design performance was sensitive to placebo patient numbers for all designs, favoring designs with a 7:1 balance in active and placebo-treated subjects (**Table 4**).

No-Go hypothetical test cases. All designs had lower frequency (14–32%) of identifying D–R correctly for the No-Go hypothetical cases than for the marketed biologic test cases (**Table 4**). Again, the success rate was ranked mehmimab (ED_{50} = 32.1 mg) > discontinumab (ED_{50} = 11.8 mg), which corresponds to the ED_{50} rank for these compounds (**Table 1**). Design performance was sensitive to placebo patient numbers for all designs, favoring designs with a 7:1 balance in active and placebo-treated subjects (**Table 4**).

Go hypothetical test cases. All designs had lower frequency (23–53%) of identifying D–R correctly for the Go hypothetical cases than for the marketed biologic test cases (**Table 4**). In contrast to the marketed and No-Go test cases, the success rate was ranked lowpomab (ED_{50} = 182 mg) > cuspmimab (ED_{50} = 32.1 mg) > nopomab (ED_{50} = 728 mg). The doses tested (21, 70, 210, and 700 mg) bracket the lowpomab and cuspmimab ED_{50}, but lower than the nopomab ED_{50}. Design performance was sensitive to placebo patient numbers for all designs, favoring designs with a 7:1 or 3:1 balance in active and placebo-treated subjects (**Table 4**).

DISCUSSION

CTS has been demonstrated in the design of Phase 2 and later trials with notable impact.[25–28] Previous uses of MBMA appear in the literature such as benchmarking a compound (post-study) against a field of competitors and powering a confirmatory trial against an expected result from an active competitor.[23,24] These two tools are combined in this work to evaluate designs for a FIP study for a class of compounds.

We have evaluated concentrated designs (subjects dosed placebo or 700 mg intravenous) and distributed designs (subjects dosed placebo or 21, 70, 210, or 700 mg) for small simulated FIP (N = 16) studies in psoriasis. Concentrated designs offered only a marginal improvement in G/NG decision making (1%) over distributed designs for the marketed biologic test cases. Concentrated designs offered improvement in G/NG decision making (3–6%) over distributed designs for the No-Go hypothetical test cases. The relative improvement in G/NG decision making for the two designs were more nuanced for the Go hypothetical cases, but the differences were nominal (3–7%). These results suggest that either design type has good G/NG decision making properties for novel compounds tested in psoriasis, and either design type can frequently identify novel compounds with good or poor marketing potential. Overall, a 1:1 and 5:3 active:placebo ratio produced the best performance for the concentrated and distributed designs, respectively.

Only the distributed designs offered information regarding D–R properties of a compound. Good D–R information was obtained 45–65, 14–32, and 23–53% for the marketed, No-Go, and Go test cases, respectively. Test cases where the correct G/NG decision was "No-Go" yielded the lowest rate of obtaining correct D–R information. However, given that a drug would be terminated at this stage of development, not having a clear understanding of the D–R is of small consequence. Overall, a 7:1 active:placebo ratio produced the best performance for the distributed designs, respectively, in contrast to the 1:1 or 5:3 active:placebo ratio favored for G/NG decision making. This analysis suggests that N = 16 subjects and the dose levels selected (21, 70, 210, and 700 mg) is not adequate to routinely (>80%) recover an ED_{50} estimate within twofold of the true value. Prospective trial design on number of subjects and/or dose selection could be used to increase the rate at which a design recovers an ED_{50} estimate within twofold of the true value.

Increasingly, small (FIP) studies in the development of biologics are conducted in patients and information on drug effect is utilized to move directly to large dose-ranging studies. Techniques established for late-stage confirmatory clinical trials provide a straightforward method for statistically testing evidence of biological activity in small patient studies. This approach is best served by studying the maximal contrast of a compound against placebo or an active comparator. However, this approach deviates from the "learning" mode of early clinical trials and places emphasis on "confirming" drug effect. Moving directly to deduction (confirmation) of efficacy is efficient, in the sense that it is rapid and of low cost. However, this bypasses a crucial induction (learning) step, i.e., learning how efficacy relates to a continuum of doses. Thus, the next learning step (Phase 2) begins with a deficiency that must somehow be corrected by addition of cost and time (more dose arms and larger enrollment, adaptive study designs, etc.), thereby decreasing efficiency. This work shows that, for a small decrease in the ability to make a correct G/NG decision, a distributed trial design provides D–R learning that can provide advantage in Phase 2.

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds *et al.*

7

A limitation of this work is that safety analysis was not included into decision making. For illustrative purposes, the focus was on efficacy, yet the principle should hold that a distributed design would be more informative for safety endpoints as well. If safety signals were detected by a distributed design with a D–R pattern, the sponsor would have guidance on an appropriate dose-selection strategy to take forward, balancing efficacy and safety. In contrast, if a safety signal was detected by a concentrated design, the sponsor would have no guidance on how lower doses may reduce the safety signal. Another potential drawback to the distributed designs is that they require analysis of response using an assumed model form. However, MBMA may provide this model, drawing on the observed D–R relationship from a number of trials to formulate this model. Finally, the marketed test cases selected in this study have observable clinical response after a single dose, but this may not be true for many other therapeutic areas that may require multiple doses and longer observation periods. However, the methodology laid out here and the implication that a distributed design can provide dose–responsive information with only small loss in G/NG power are also compatible with small, multiple-dose studies in indications for which longer observation periods are required.

In this work, a clinically relevant endpoint (percent change in PASI for psoriasis patients) was selected as an illustration. Other endpoints that are not "registrational" in nature (e.g., biomarkers, clinical signs, etc.) are amenable to this type of analysis. However, when a "registrational" endpoint is available, our instincts pull us towards traditional, late-phase confirmatory analyses to demonstrate that the drug is different from placebo. Our instincts, examined under the conditions in this paper, are not supported: concentrated designs are not substantially better for demonstrating a drug "works" than distributed designs in this context. Moreover, distributed designs offer information regarding the D–R properties of a compound, which facilitates better design of the pivotal dose-ranging Phase 2 study.

METHODS

MBMA of biologic treatments for psoriasis. A MBMA was performed of publically available data from randomized controlled trials of approved compounds in psoriasis and psoriatic rheumatoid arthritis. Data to inform the meta-analysis were found via literature search (PubMed) and review of reference lists from previous meta-analyses, clinicaltrials.gov, conference abstracts, and corporate websites. The included data represented 27 trials of antitumor necrosis factors, ustekinumab, and methotrexate. The MBMA proceeded as joint modeling of the probability of achieving PASI 50, 75, 90, and mean percent improvement in PASI score. Placebo and treatment effect were modeled for each drug. The placebo response was assumed to be different for each trial and the treatment effect was modeled using an E_{max} model with a different E_{max} for each endpoint and drug class. We evaluated impact of endpoint, drug, drug class, regimen, indication (psoriasis vs. psoriatic rheumatoid arthritis), failure of prior treatment, baseline PASI score, disease duration, age, weight, and gender on model parameters. For this paper, the model for percent improvement in PASI score was used.

Clinical trial simulation. To bound the percent improvement in PASI score [$-\infty\%$, 100%] in the real domain, the ratio of on treatment PASI score ($PASI_{treatment}$) over baseline PASI score ($PASI_{base}$) was simulated in the log domain according to the following equation:

$$LPASI = \log\left(\frac{PASI_{treatment}}{PASI_{base}}\right) = E_0 + \frac{E_{max} \cdot Dose}{ED_{50} + Dose} \quad (1)$$

where E_0 reflects the placebo response; E_{max}: the maximal drug response; and ED_{50}: the dose providing half maximal drug response. Please note that the parameters E_0, E_{max}, and ED_{50} reflect the response in the log domain. A similar transformation was used for the meta-analysis.

Variability at the subject level was added using a linear model (slope = sERR, intercept = iERR) of SD, dependent on LPASI:

$$Y = LPASI + (sERR*LPASI+iERR)*\varepsilon \quad (2)$$

$$\varepsilon \sim N(0,1) \quad sERR = -0.32 \quad iERR = 0.30 \quad (3)$$

where Y is the subject-level PASI response in the log domain. Note that sERR is given with a negative sign, as LPASI is negative. Only one PASI observation was simulated per subject. The parameters of the between-subject variance model were selected to recover the between-subject variability observed in the PHOENIX 1 and 2 trials for ustekinumab.[21] Mean and between-subject variability percent PASI improvement (point: mean, lines: ±SD) for placebo, 45-mg ustekinumab, and 90-mg ustekinumab from the simulations (red) and PHOENIX2 data (blue) are compared in **Figure 3**. Note that
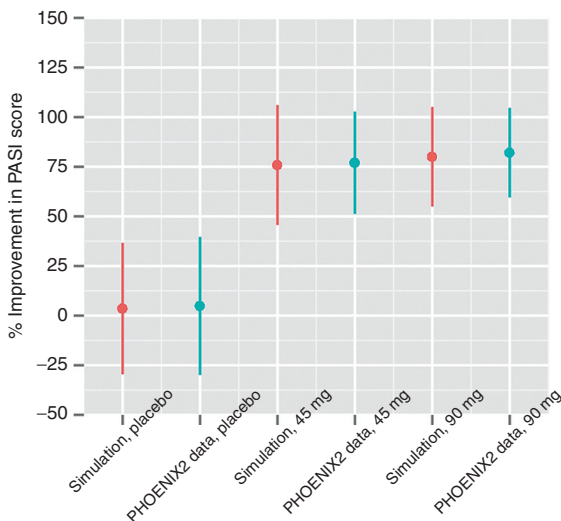
Figure 3 Arithmetic mean and between-subject variability percent PASI improvement (point: mean, lines: ±SD) for placebo, 45-mg ustekinumab, and 90-mg ustekinumab from the simulations (red) and PHOENIX2 data (blue).[21]

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds *et al.*

8

the between-subject variability decreases with increasing response (but not as much as expected with a simple additive error model in the log domain), which motivated this choice of variance model.

Trials, ($N = 9,999$; the maximum allowed using NONMEM 7.2, ref. 29) for each of the eight designs and eight drugs were simulated ($N = 16$ subjects per trial) using Eqs. 1–3 using NONMEM 7.2 on a cluster. Each design was given as a single control stream, each containing the eight drugs, and processed using message passing interface parallelization on six cores (8 designs × 6 cores/design = 48 cores). Total simulation and back-estimation time was less than 2 days.

For each simulated trial, estimates of population typical E_{max}, E_0, ED_{50}, sERR, and iERR were obtained using an estimation model that varied by design type and subject dose assignment:

$$\text{LPASI} = \begin{cases} \text{Concentrated Design} & E_0 + E_{max} \\[2ex] \text{Distributed Design} & E_0 + \dfrac{E_{max} \cdot \text{Dose}}{ED_{50} + \text{Dose}} \end{cases} \quad (4)$$

Thus, the drug effect (E_{drug}) at the maximum feasible dose (700 mg) is derived by observation (concentrated designs) or estimated (distributed designs):

$$E_{drug} = \begin{cases} \text{Concentrated Design} & E_0 + E_{max} \\[2ex] \text{Distributed Design} & E_0 + \dfrac{E_{max} \cdot 700}{ED_{50} + 700} \end{cases} \quad (5)$$

Given the log transformation, $e^{E_{drug}}$ is the geometric mean ratio of on treatment PASI over baseline PASI and $1 - e^{E_{drug}}$ is the PASI % improvement from baseline, including placebo response, at the maximum feasible dose. The G/NG criteria for each simulated trial were calculated after back-transformation of E_{drug} and E_0 to the original scale for PASI percent change from baseline and then cast as a binary outcome by comparison to the G/NG criteria of a 50% absolute difference from placebo:

$$\text{G/NG} = \left[\left[1 - e^{E_{drug}}\right] - \left[1 - e^{E_0}\right] > 0.50 \right] \quad (6)$$

The G/NG criteria were contrasted with the correct G/NG decision at the true values of E_{drug} and E_0 used in the simulation.

The rate at which a design resulted in a correct G/NG decision was then computed by summing the number of correct G/NG decisions and dividing by the number of trials (9,999).

Similarly, the "D–R" criteria for each simulated trial were calculated and contrasted with the true value (ED^*_{50}) used in the simulation, after back-transformation to the real domain:

$$\text{Good D–R} = \left[1/2 \le \frac{ED_{50}}{ED^*_{50}} \le 2 \right] \quad (7)$$

The rate at which a design resulted in good D–R information was then computed by summing the number of good D–R and dividing by the number of trials (9,999).

See **Supplementary Methods** online for model specification and estimation steps.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

✓ Decisions to move forward into Phase 2 (G/NG) are increasingly made by small studies in relevant patient populations that are designed to demonstrate biologic impact. A common trial design employed to address this question contrasts patient response at the highest feasible dose and placebo treatment.

WHAT QUESTION THIS STUDY ADDRESSED?

✓ Does a dose-ranging study design offer similar G/NG decision-making quality with the additional benefit of providing valuable D–R information needed to inform Phase 2 dose selection?

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

✓ For a minor decrease in the ability to make a correct G/NG decision, a dose-ranging trial design provides valuable D–R information needed to guide Phase 2 dose selection.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

✓ CTS can increase the efficiency and quality of drug development decision making by studying the limitations and benefits of study designs prospectively.

1. Administration, U.F.A.D. *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products.* <http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf> (2004).
2. Lalonde, R.L. *et al.* Model-based drug development. *Clin. Pharmacol. Ther.* **82**, 21–32 (2007).

CTS to Inform Phase 2: Concentrated vs. Distributed FIP Designs in Psoriasis
Dodds *et al.*

npg

9

3. Grasela, T.H. & Slusser, R. Improving productivity with model-based drug development: an enterprise perspective. *Clin. Pharmacol. Ther.* **88**, 263–268 (2010).
4. Allerheiligen, S.R. Next-generation model-based drug discovery and development: quantitative and systems pharmacology. *Clin. Pharmacol. Ther.* **88**, 135–137 (2010).
5. Reigner, B.G., Williams, P.E., Patel, I.H., Steimer, J.L., Peck, C. & van Brummelen, P. An evaluation of the integration of pharmacokinetic and pharmacodynamic principles in clinical drug development. Experience within Hoffmann La Roche. *Clin. Pharmacokinet.* **33**, 142–152 (1997).
6. Aarons, L., Karlsson, M.O., Mentré, F., Rombout, F., Steimer, J.L. & van Peer, A.; COST B15 Experts. Role of modelling and simulation in Phase I drug development. *Eur. J. Pharm. Sci.* **13**, 115–122 (2001).
7. Mandema, J.W. *et al.* Model-based development of gemcabene, a new lipid-altering agent. *AAPS J.* **7**, E513–E522 (2005).
8. Mandema, J.W., Cox, E. & Alderman, J. Therapeutic benefit of eletriptan compared to sumatriptan for the acute relief of migraine pain–results of a model-based meta-analysis that accounts for encapsulation. *Cephalalgia* **25**, 715–725 (2005).
9. Hing, J., Perez-Ruixo, J.J., Stuyckens, K., Soto-Matos, A., Lopez-Lazaro, L. & Zannikos, P. Mechanism-based pharmacokinetic/pharmacodynamic meta-analysis of trabectedin (ET-743, Yondelis) induced neutropenia. *Clin. Pharmacol. Ther.* **83**, 130–143 (2008).
10. Luu, K.T., Raber, S.R., Nickens, D.J. & Vicini, P. A model-based meta-analysis of the effect of latanoprost chronotherapy on the circadian intraocular pressure of patients with glaucoma or ocular hypertension. *Clin. Pharmacol. Ther.* **87**, 421–425 (2010).
11. 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).
12. Gibbs, J.P. *et al.* Quantitative model of the relationship between dipeptidyl peptidase-4 (DPP-4) inhibition and response: meta-analysis of alogliptin, saxagliptin, sitagliptin, and vildagliptin efficacy results. *J. Clin. Pharmacol.* **52**, 1494–1505 (2012).
13. Rohatagi, S. *et al.* Model-based development of a PPARγ agonist, rivoglitazone, to aid dose selection and optimize clinical trial designs. *J. Clin. Pharmacol.* **48**, 1420–1429 (2008).
14. Salinger, D.H., Endres, C.J., Martin, D.A. & Gibbs M.A. A semi-mechanistic model to characterize the pharmacokinetics and pharmacodynamics of brodalumab in healthy volunteers and subjects with psoriasis in a first-in-human single ascending dose study. *Clin. Pharmacol. Drug Dev.* (in review).
15. Ng, C.M., Joshi, A., Dedrick, R.L., Garovoy, M.R. & Bauer, R.J. Pharmacokinetic-pharmacodynamic-efficacy analysis of efalizumab in patients with moderate to severe psoriasis. *Pharm. Res.* **22**, 1088–1100 (2005).
16. Ashcroft, D.M., Wan Po, A.L., Williams, H.C. & Griffiths, C.E. Clinical measures of disease severity and outcome in psoriasis: a critical appraisal of their quality. *Br. J. Dermatol.* **141**, 185–191 (1999).
17. Gottlieb, A.B. *et al.* A phase 1, double-blind, placebo-controlled study evaluating single subcutaneous administrations of a human interleukin-12/23 monoclonal antibody in subjects with plaque psoriasis. *Curr. Med. Res. Opin.* **23**, 1081–1092 (2007).
18. Holford, N., Ma, S.C. & Ploeger, B.A. Clinical trial simulation: a review. *Clin. Pharmacol. Ther.* **88**, 166–182 (2010).
19. Menter, A. *et al.* Adalimumab therapy for moderate to severe psoriasis: A randomized, controlled phase III trial. *J. Am. Acad. Dermatol.* **58**, 106–115 (2008).
20. Leonardi, C.L. *et al.*; PHOENIX 1 study investigators. Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 76-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 1). *Lancet* **371**, 1665–1674 (2008).
21. Papp, K.A. *et al.*; PHOENIX 2 study investigators. Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). *Lancet* **371**, 1675–1684 (2008).
22. Kavanaugh, A. *et al.* Golimumab, a new human tumor necrosis factor α antibody, administered every four weeks as a subcutaneous injection in psoriatic arthritis: Twenty-four-week efficacy and safety results of a randomized, placebo-controlled study. *Arthritis Rheum.* **60**, 976–986 (2009).
23. Mandema, J.W., Salinger, D.H., Baumgartner, S.W. & Gibbs, M.A. A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis. *Clin. Pharmacol. Ther.* **90**, 828–835 (2011).
24. Mandema, J.W., Gibbs, M., Boyd, R.A., Wada, D.R. & Pfister, M. Model-based meta-analysis for comparative efficacy and safety: application in drug development and beyond. *Clin. Pharmacol. Ther.* **90**, 766–769 (2011).
25. Miller, R. *et al.* How modeling and simulation have enhanced decision making in new drug development. *J. Pharmacokinet. Pharmacodyn.* **32**, 185–197 (2005).
26. Bonate, P.L. Clinical trial simulation in drug development. *Pharm. Res.* **17**, 252–256 (2000).
27. 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).
28. Lockwood, P.A., Cook, J.A., Ewy, W.E. & Mandema, J.W. The use of clinical trial simulation to support dose selection: application to development of a new treatment for chronic neuropathic pain. *Pharm. Res.* **20**, 1752–1759 (2003).
29. NONMEM user's guide (1989–2009) [computer program] (ICON Development Solutions, Ellicott City, MD, 2009).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.nature.com/psp)