nature publishing group

Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development

PA Milligan[1], MJ Brown[2], B Marchant[3,10], SW Martin[1], PH van der Graaf[4,1], N Benson[4,11], G Nucci[5], DJ Nichols[5], RA Boyd[6], JW Mandema[7], S Krishnaswami[6], S Zwillich[8], D Gruben[2], RJ Anziano[2], TC Stock[9] and RL Lalonde[6]

The pharmaceutical industry continues to face significant challenges. Very few compounds that enter development reach the marketplace, and the investment required for each success can surpass $1.8 billion. Despite attempts to improve efficiency and increase productivity, total investment continues to rise whereas the output of new medicines declines. With costs increasing exponentially through each development phase, it is failure in phase II and phase III that is most wasteful. In today's development paradigm, late-stage failure is principally a result of insufficient efficacy. This is manifested as either a failure to differentiate sufficiently from placebo (shown for both novel and precedented mechanisms) or a failure to demonstrate sufficient differentiation from existing compounds. Set in this context, this article will discuss the role model-based drug development (MBDD) approaches can and do play in accelerating and optimizing compound development strategies through a series of illustrative examples.

The first article in this series[1] outlined the rationale for model-based drug development (MBDD) and introduced six key components (pharmacokinetics–pharmacodynamics (PK–PD) and disease models, competitor information and meta-analysis, design and trial execution models, data analysis models, quantitative decision criteria, and trial performance metrics) of a "quantitative toolkit." Examples were used to illustrate their application across a range of drug-development contexts. This article examines and illustrates the continued evolution of MBDD since the earlier publication, with particular emphasis on the requirements to implement and maintain application of this concept within a large multinational pharmaceutical research and development organization.

The evolution of MBDD is depicted in **Figure 1**. Going beyond solely development-related activities, it also now encompasses (earlier) discovery and (later) real-world clinical utilization settings. **Figure 1** has many familiar elements, such as visualizing the entire drug discovery and development process through to commercialization and beyond as a series of "left to right" events/stages. However, we also emphasize the need for both the feed-forward and feed-back of information across all these stages. This requires a much greater degree of both data and (inferential) knowledge integration than is currently the norm. This requirement represents a significant technical challenge given the complexity of the drug discovery and development process. The linear nature of the current process creates an opportunity for the variety of quantitative approaches encompassed in **Figure 1** to accelerate drug development. This opportunity will be realized through not treating the overall process as a series of discrete events that occur in a sequential manner in which successful completion of a particular stage enables compound continuation to the adjoining stage. MBDD offers the potential to treat the overall process as a continuum of integrated and interrelated events that occur in an order that can be adjusted in light of the most complete appreciation of a compound's inherent risks informed by the capacity of the accumulated data and (inferential) knowledge integration to mitigate and manage these risks.

In **Figure 1**, there are not only (horizontal) "left to right" elements to consider but also a series of (vertical) "top to bottom" elements, which are important. We regard each of these

[1]Global Clinical Pharmacology, Pfizer, Sandwich, UK; [2]Department of Statistics, Pfizer, Groton, USA; [3]Department of Clinical Development, Pfizer, Sandwich, UK; [4]Department of Pharmacokinetics Dynamics and Metabolism, Pfizer, Sandwich, UK; [5]Global Clinical Pharmacology, Pfizer, Cambridge, Massachusetts, USA; [6]Global Clinical Pharmacology, Pfizer, Groton, Connecticut, USA; [7]Quantitative Solutions, Inc., Menlo Park, California, USA; [8]Department of Clinical Development, Pfizer, Groton, Connecticut, USA; [9]Department of Clinical Development, Pfizer, Collegeville, Philadelphia, USA; current affiliations: [10]Marchant Biopharm Consulting Ltd, Kowloon, Hong Kong; [11]Xenologiq, Denne Hill Business Park, Canterbury, UK. Correspondence: PA Milligan (peter.a.milligan@pfizer.com)

Received 27 December 2012; accepted 4 March 2013; advance online publication 1 May 2013. doi:10.1038/clpt.2013.54

components as foundational elements that are instrumental in determining the likelihood of a given compound ever being prescribed to patients. The traditional paradigm described previously (historical enabling of activities rather than informing of future activities) often results in a situation in which emerging weakness in earlier foundational elements (rationale) forces organizations into the very difficult situation of having to adopt later "shoring up" activities, which come at great cost with often little return in terms of increasing the compounds' probability of success. This challenging cycle of events has impacted the level of productivity of the industry as a whole. Once again, we see an opportunity for the variety of quantitative approaches encompassed in **Figure 1** to accelerate drug development. We can continuously examine each single and combined confidence in rationale (selected pathway, target, molecule, dose regimen, patients) in light of emerging data, both internal and external, and therefore more reliably inform our subsequent go/no go decisions.

MBDD IMPLEMENTATION PERSPECTIVES

Our organization embarked on an enterprise-wide adoption of MBDD concepts in 2005 in light of what was considered to be an unsustainable (albeit industry average) late-stage study failure rate, with failure being defined as the inability of the emerging study efficacy data to correspond to the study's *a priori* efficacy hypothesis. We examined the causes of such failures in 68 historical phase II/III/IV clinical trials and found a number of recurring themes, which were considered to be "root causes" and necessitate improvement:

1. Insufficient characterization of the exposure–response relationship before implementing confirmatory studies in late-stage clinical development;
2. Insufficient knowledge of the treatment effect in the target population (difference from placebo or active comparator and/or variance);
3. Incomplete knowledge of the drug, mechanistically related drugs, and attributes of the therapeutic indication of interest because relevant data were not systematically collated, stored, and utilized;
4. Lack of team experience with the primary end point (often due to "enhancements" of the historically established end point).

These findings demonstrated both technical and behavioral/cultural inadequacies. Consequently, attempts to reverse this situation had to rectify both of these dimensions in a staged manner. We were aware, through observing the range of organizational structures and functions in existence at that time, that organizational constructs could significantly limit or enhance adoption of any corrective approaches. This organizational heterogeneity existed despite a high degree of concordance across many pharmaceutical companies in their stated desire to increase both the utilization and influence of MBDD approaches.

We initially observed that the expertise required to deliver the technical components of MBDD resided within discrete and dislocated groupings; longitudinal MBDD within clinical pharmacology/pharmacometrics, study design within statistics, and program strategy and design within the therapeutic area clinicians. A second observation was that none of the technical components of MBDD envisioned at that time were in fact truly innovative. We were facing a situation in which the desired methodologies were established, and we had some colleagues with the skills to implement them, but we had an organization that did not fully capitalize on the available opportunity. Therefore, to shift the dynamic from relying solely on the more technically able colleagues who existed across the key disciplines (clinicians, statisticians, and clinical pharmacologists/pharmacometricians) driving organizational implementation in a disaggregated "bottom–up" manner, we sought and received considerable and active senior-level sponsorship before and during implementation in a "top–down" manner. This senior-level support was manifested in a number of ways, including repeated and unambiguous public statements both internally and externally on the importance of MBDD; incorporation of MBDD into organizational and individual goals; and the alignment of governance committees, technical review committees, and drug-development teams on their specific responsibilities to "institutionalize" MBDD within the organization.

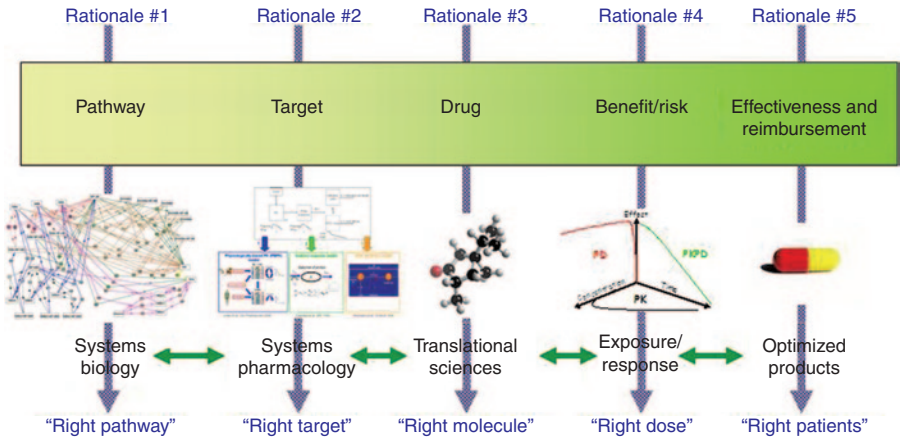

Figure 1 The evolution of model-based drug development (MBDD). Adapted from ref. 2.

Two types of educational offerings were developed and delivered. First, colleagues within each discipline were trained on what was expected of them from a technical perspective and what they needed to understand of the other two disciplines in order to be able to interact effectively. Second, the three disciplines were brought together to participate in a series of case-based workshops with the emphasis on working together to (overtly) solve drug-development problems synergistically and to (covertly) exemplify the benefits of improved team behaviors. Over a 6-month period, 90% of eligible colleagues across all countries where clinical development activities occurred completed the required training.

Participants were trained to recognize that as the types of development questions differed, the strategies and study designs suited to address these questions should also differ. The technical aspects of our training program followed the learning/confirming cycle concept introduced by Lewis Sheiner in 1997.[3] Earlier studies were designed to maximize learning opportunities. The resulting optimized study data, positioned within the relevant compound landscape, informed the subsequent strategy decisions. It was demonstrated that learning study designs used analysis methods that would not necessarily be the best choice for confirmatory studies. For example, teams could use models based on reasonable assumptions that emanated from the broadest knowledge base instead of using the minimal assumption analyses typically required for key regulatory decisions.

Participants were also trained to recognize that MBDD could and should be used in the design phase to help assess design properties, beyond typical sample size methods that focus on false-positive and false-negative rates. It was important for the participants to understand the operating characteristics of a design (e.g., probability of making a correct decision, precision of estimates, size or cost of the study, feasibility of accrual, and time to achieve an answer) under various assumptions and the ability of a particular study to meet the objectives of the program. These operating characteristics could be determined using direct computation with assumed underlying probability distributions (typical for standard power/sample size estimates but also available for other characteristics) or be simulation based following the generation of data intended to mimic the proposed trial and analysis. These data could be generated based on hypothetical population characteristics or using observed data from other sources (or some combination). A range of designs is available for consideration, and over the years following MBDD implementation we have seen an increase in the adoption of various flexible/adaptive designs such as the classic group sequential approach that includes early termination for success or futility, and the more innovative adaptive designs that include termination of some dose groups based on interim efficacy or safety results or alteration of the randomization ratios to increase precision in the determination of an effective dose range.

What was continuously emphasized throughout the training was that MBDD was not solely about applying these designs; instead, it encouraged teams to consider design options within a "fit for purpose" context. For example, if the intent was to understand and characterize a dose– or exposure–response relationship, then a study should be designed to optimize the amount of information required to meet that objective. Previously in phase II, confirmatory-type designs, in which each dose was compared with placebo (or comparator), were conducted, based on the argument that these studies (if successful) may qualify as one of the confirmatory studies required for regulatory approval. However, given the modest success rate of such studies in our historical analysis and the limited learning derived from these designs, this is not an efficient strategy. We showed that a study designed to enable estimation of an exposure–response relationship (given a particular structural form) could be much smaller, yet yield much more (precise) information, thereby facilitating dose selection decisions. We also showed that the resultant information would be extremely important in supporting the ultimate approval of the compound because the exposure–response relationship had been characterized in a "well-controlled" study and subsequently informed the proposed dose(s) justification. Later, phase III confirmatory study designs could then be informed using a more comprehensive efficacy "evidence base" that maximizes the probability of success for both the studies and the compound. In our experience, this particular change (from hypothesis testing toward estimation approaches as the primary data analytic) was the single largest factor in reducing our overall clinical trial budget. **Table 1** illustrates the nature and extent of the efficiencies gained through application of MBDD across a range of indications.

In the years following MBDD implementation, our organization experienced an improvement in late-stage clinical development productivity (**Figure 2**). Of greatest importance was the incremental gain in the proportion of successful phase III and IV trials. We attribute this success to a number of factors, including study teams developing comprehensive knowledge management strategies that enabled a more accurate quantification of the probability of achieving the required product profile (which was regularly updated as data accumulated throughout compound development). This approach enabled more robust decision making because study teams avoided performing studies that had an unacceptable probability of failing. As a consequence, more compounds were terminated earlier in phase II; however, it should be recognized that the cost of many of these failures was considerably reduced through early termination for futility. Earlier termination also means that resources could be reallocated to other compounds in development that have a greater probability of success.

The proportion of successful phase II trials was relatively static year on year. There is potentially a complex dynamic in existence in phase II; on the one hand, there are factors that increase the phase II study failure rate (unacceptable probability of success in phase III); and on the other, there is the potential for some of the approaches shown in **Table 1** to minimize and mitigate phase II failure rates. The interaction of these various complexities could account for the results shown in **Figure 2** and although achieving phase II stasis is a "good" result, we envision significant challenges to increase

the proportion of successful phase II trials, which are discussed later in this article.

As mentioned previously, efficiency gains could be achieved via the nature/attributes of the actual studies performed and the risk/cost mitigation strategy for the program as a whole. In the first year of implementation, we closely tracked the cost efficiencies gained through designing MBDD "informed" studies over our historical norms for the specific indication, and we were able to document returns of just under $70 million. Following 2 years of implementation, once MBDD principles were firmly established within the organization, $100 million was taken out of the global clinical trial budget, year on year. This encouraged teams to use their enhanced technical and behavioral attributes to "do the same for less" or in some instances "do more for less." This has been achieved through what has largely become the organizational norm: to quantitatively articulate protocol and program risks with a range of buy-up/buy-down options enumerated for discussion and selection.

The examples that follow illustrate the evolution of MBDD application over the past few years within our organization.

Application of integrated pharmacometrics and systems pharmacology models to guide decision criteria and clinical study designs for early clinical development

Endometriosis is a gynecological condition that affects up to 15% of women of reproductive age and that results from the presence of endometrial-like tissue outside of the uterine cavity and requires estrogen for proliferation.[4] The major clinical symptoms are dysmenorrhea (pain during menses), pelvic pain with lower-back or abdominal pain, dyspareunia (pain during sex), and dysuria (urinary urgency, frequency, and painful voiding). Treatments that are effective at reducing endometriosis-related pain include the use of gonadotropin-releasing

Table 1 Some examples of the high-level efficiencies gained over historical designs and data analytics following MBDD implementation

Indication	MBDD approach adopted	Efficiencies gained over historical designs and analysis
Thromboembolism[a]	Omit phase IIa, model-based dose–response relationship, adaptive phase IIb design	2,750 Fewer patients, 1 year shorter study duration
Hot flashes	Model-based dose–response relationship	1,000 Fewer patients
Fibromyalgia	Prior data supplementation, model-based dose–response relationship, sequential design	760 Fewer patients, 1 year shorter study duration
Type 2 diabetes	Prior data supplementation, model-based dose–response relationship	120 Fewer patients, 1 year shorter study duration
Gastroesophageal reflux	Model-based dose–response relationship	1,025 Fewer patients
Rheumatoid arthritis	Model-based dose–response relationship	437 Fewer patients, increased probability of success
Global anxiety disorder	Omit phase IIb	260 Fewer patients, 1 year shorter study duration
Lower urinary tract symptoms	Meta-analysis	Increased probability of success
Urinary incontinence	Meta-analysis	Increased probability of success

MBDD, model-based drug development.

[a]This application is discussed further in the text as example 4, "Adaptive dose-finding phase II study designed using clinical trial simulations."

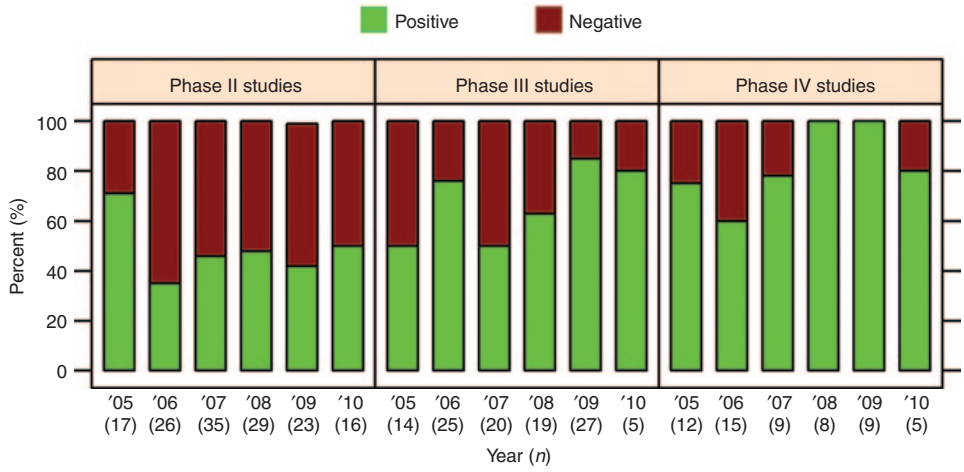

Figure 2 Phase II, III, and IV study outcomes (% positive or negative for the primary efficacy outcome) following MBDD implementation. The year refers to when each specific study was initiated. The number in parenthesis provides the actual number of studies to subsequently complete and form the basis of the reported percentages. MBDD, model-based drug development.

hormone (GnRH) analogs.[5] However, these compounds cause profound decreases in estrogen and as a result, patients often have menopause-like side effects such as hot flashes and reduction in bone mineral density (BMD). Although GnRH analogs provide very effective pain treatment, the side effects limit their use to 6 months. Low-dose estrogen–progestin hormone therapy has been used to prevent bone loss when prolonged treatment is needed.[6–9] We had identified a number of potential targets in the GnRH pathway as treatments for endometriosis. An 8-week proof-of-concept (POC) study for endometriosis was being considered, but it would not be long enough to observe the key safety issue of BMD loss.

A number of key clinical development questions were therefore identified, which include the following:

1. What is the optimal range of estrogen levels?
2. Can modulation of the GnRH pathway achieve ideal estrogen levels?
3. Which biomarkers (e.g., estrogen and bone markers), if any, would provide reliable predictions of long-term BMD changes?
4. Can an optimal biomarker range be identified?
5. What is the expected biomarker time course?

Due to the highly nonlinear nature and multiple feedback loops in the female menstrual cycle and bone metabolism, an integrated pharmacometrics and systems pharmacology model-based approach was undertaken to address these questions.[10]

A standard pharmacometrics approach was used to determine the relationship between estrogen (estradiol; E2) levels and endometrium symptoms index score (ESSS). Internal patient-level (n = 499) data (E2 and ESSS measurements, n = 1,354) were obtained from three clinical studies with nafarelin, a marketed GnRH agonist.[11–13] A logistic regression model was developed for ESSS in which a cumulative logit function, including E2 levels for each patient and visit and incorporating interindividual variance, determines probabilities for each ESSS score from 0 to 3 or ≥4. The logistic regression results characterized the E2–ESSS relationship in which the probability of a patient experiencing a less severe score decreased as E2 levels increased (**Figure 3**).

The relationship among bone markers, lumbar spine BMD, and E2 was determined by extension of an existing multiscale model of calcium and bone homeostasis.[13] To predict longitudinal effects of E2 suppression on lumbar BMD, data were extracted from 14 double-blind studies reporting lumbar spine BMD and E2 following treatment (>6 months) with GnRH agonists or antagonists for endometriosis. The effects of estrogen on several model components, including transforming growth factor-β, osteoblasts, active osteoclasts, tubular reabsorption, and renal excretion of calcium, were simultaneously fitted to the clinical data. The extended multiscale model was then used to predict the effects of GnRH-related alterations on the bone formation marker, bone-specific alkaline phosphatase, and the degradation markers, *N*-telopeptide or *C*-telopeptide. The effect of GnRH modulation was modeled through reductions in E2 levels

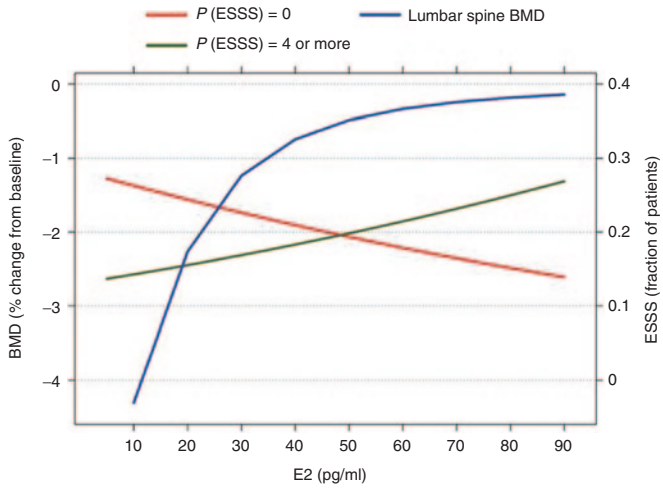

Figure 3 Comparison of predicted endometrial symptoms severity score (ESSS) probabilities and lumbar spine bone mineral density (BMD) change as a function of serum estradiol (E2). ESSS model predictions represent the median probabilities from the ordered categorical logistic regression model; BMD predictions were taken from the deterministic multiscale systems pharmacology model.

ranging from 90 to 60%. The estimate of lumbar spine BMD changes was modeled via changes in osteoblast (bone-specific alkaline phosphatase) and osteoclast (*C*-telopeptide) function. The time course of bone biomarker changes helped identify which, if any, of the commonly available bone biomarkers would offer an early, sensitive measure of long-term BMD changes as well as define the target range for the biomarker response. These models indicated that marked reductions in estrogen levels (80%) were predicted to cause minor changes in biomarker levels (<25%) and very minor changes in BMD (<1%) during the first 3 months following treatment. Six months of continuous suppression of estrogen (80%) was predicted to cause a 2% BMD loss, which is twice the rate following menopause.

An acceptable BMD loss of ≤1% could be achieved with ~60% suppression of estrogen levels and was accompanied by minimal bone marker changes (<10%). Therefore, in the dose range of interest (60–80% inhibition of estrogen), biomarkers of bone turnover are unlikely to be able to adequately differentiate between doses. By contrast, estrogen level appeared to be the most sensitive biomarker for both efficacy and safety. Maximal suppression of estrogen was predicted to occur within 1–2 months of treatment and therefore is a rapid and sensitive biomarker of treatment effect. A target range of estrogen levels of 20–40 pg/ml was identified to provide a clinically meaningful improvement on ESSS pain score, with limited effects on 6-month BMD (<2% change; **Figure 3**). In addition to the multiscale systems pharmacology model of calcium homeostasis, a systems pharmacology model of the female menstrual cycle was developed to assess the feasibility of achieving the desired target range of estrogen suppression via modulation of the GnRH pathway, as well as supporting early clinical development programs.[14] This model highlighted the difficulties in being able to target a narrow range of estrogen levels in a diverse patient population.

In summary, this work identified target levels for estrogen that would provide symptomatic pain relief with minimal impact on BMD. A systems pharmacology model of the female menstrual cycle[14] indicated that targeting the GnRH pathway to achieve the desired range of serum estrogen levels would be difficult to achieve; therefore, the research program was halted before any compound entered the clinic.

Application of systems pharmacology modeling to the development of fatty acid amide hydrolase inhibitors for pain

The endogenous cannabinoid (endocannabinoid) *N*-arachidonoyl ethanolamine (anandamide, AEA) activates the G-protein-coupled receptors CB1 and CB2, which are believed to play a role in nociceptive signaling. CB1 and CB2 receptors are the primary target of the main psychoactive substance, Δ9-tetrahydrocannabinol, in marijuana, but the potential analgesic clinical benefits of Δ9-tetrahydrocannabinol have been difficult to demonstrate due to its well-known psychoactive actions and other side effects.[15] Fatty acid amide hydrolase (FAAH) is an integral membrane enzyme that is involved in degradation of AEA and related fatty acid amides, including linoleoyl ethanolamide, *N*-palmitoylethanol amine, oleoylethanolamide, and stearoylethanolamide. It has therefore been postulated that inhibition of FAAH could provide a novel pharmacological strategy to enhance the "endocannabinoid tone," and preclinical studies have provided support for the hypothesis that inhibition of FAAH can be associated with analgesic effects in certain rodent models.[16] Therefore, several groups have been pursuing the development of selective FAAH inhibitors for the treatment of a variety of indications, including pain.[17] We identified a potent, irreversible and selective FAAH inhibitor, PF-04457845.[18] This compound demonstrated good central nervous system penetration[19] and efficacy in preclinical models of inflammatory and noninflammatory pain,[20] although the effects were not seen consistently across a wide range of models.[16] PF-04457845 was well tolerated in first-in-human studies and displayed good PK properties, supporting further clinical development.[21] *Ex vivo* analysis in isolated leukocytes demonstrated that PF-04457845 produced near-maximal inhibition of FAAH activity in healthy subjects following oral dosing and, of note, mean fatty acid amide concentrations (including AEA) were increased 3.5- to 10-fold following PF-04457845,[21] demonstrating target engagement and modulation according to the "three pillars" principles[22] (exposure at the target site of action, binding to target, and expression of pharmacology).

During this stage of preclinical-to-clinical transition and early clinical research, it was decided to develop an integrated systems pharmacology model of the FAAH/CB1 pathway as a quantitative framework for interpretation of the emerging biomarker data. The systems pharmacology approach integrated PK–PD, physiologically based PK, and systems biology.[23] The model integrated literature and in-house data on the metabolism of AEA, palmitoylethanol amine, oleoylethanolamide, and linoleoyl ethanolamide; the kinetics of the enzymes controlling endocannabinoid substrate production and degradation; and CB1 receptor binding kinetics. Overall, the model comprised four compartments (brain, blood–brain barrier, plasma, and rest of the body), 77 reactions, and 146 parameters (see **Supplementary Figure S1** online). The key outcome of this effort was that it was not possible to replicate the biomarker profile obtained in the healthy volunteer study[21] when it was assumed that AEA was metabolized only by FAAH. The experimental clinical data showed that AEA levels increased to a plateau following administration of PF-04457845 and that increasing doses affected the width, but not the height of the plateau (**Figure 4a**). Such profiles could be replicated by the model only when it was assumed that there was another, FAAH-independent, route of metabolism involved in AEA breakdown (**Figure 4b**). Of the numerous candidates (including cyclooxygenases and cytochromes P450), further model simulations led to the hypothesis that this additional clearance process was most likely due, at least in part, to the enzyme *N*-acylethanolamine hydrolyzing acid amidase. Validation of this hypothesis requires further work.

A further conclusion derived from the systems pharmacology model was that as a result of the presence of the additional metabolism process, the maximum increases in AEA levels in the central nervous system were predicted to be relatively small and saturable within the range of PF-04457845 doses tested. AEA occupancy at the CB1 receptor was predicted to increase from 3% in the absence of treatment to a maximum of only ~25% in the presence of PF-04457845 (**Figure 4c**), unless it was assumed that AEA levels at the site of action were to be enhanced as a result of partitioning in lipids surrounding the CB1 receptor. These predictions triggered discussions that highlighted the significant gap in the project team's understanding of the quantitative relationship among AEA levels, expression of CB1 pharmacology, modulation of nociceptive signaling, and analgesic efficacy. PF-04457845 was subsequently tested for analgesic effects in patients with osteoarthritis but was, in contrast to naproxen, not differentiated from placebo.[24] Various possible explanations were provided for this lack of clinical efficacy and the apparent disconnect with promising preclinical data.[15,24] However, the systems pharmacology model provided a novel argument that inhibition of FAAH alone may lack the intrinsic "horsepower" to sufficiently modulate the endocannabinoid system and that combined inhibition of multiple AEA-metabolizing enzymes (including *N*-acylethanolamine hydrolyzing acid amidase) may be required to achieve analgesic efficacy. The quantitative systems pharmacology approach also highlighted the fact that robust methods to quantify AEA CB1 receptor occupancy and activation in a clinical setting need to be developed first to enable a more rational development of novel drugs in this field.

Overall, this case study illustrates the potential value of prospective, integrated systems pharmacology modeling and simulation for the selection and validation of drug targets. This conclusion is supported by results emerging from similar systems pharmacology approaches in pain research (e.g., focusing on other key pathways involved in the regulation of pain, such as nerve growth factor),[23,25] as well as other areas such as central nervous system diseases,[26] oncology,[27] osteoporosis,[13] endometriosis,[10] and safety.[28]

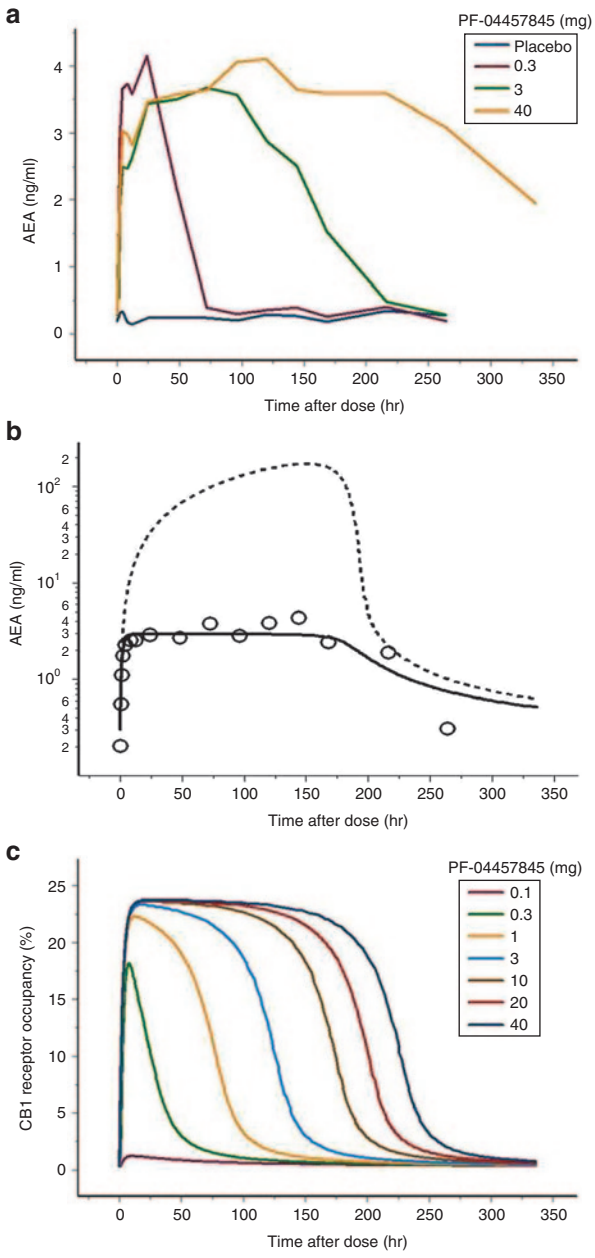

Figure 4 Observed data and systems pharmacology model predictions for the effects on anandamide (AEA) of the selective fatty acid amide hydrolase (FAAH) inhibitor, PF-04457845. (**a**) Elevation of AEA observed in a typical healthy volunteer subject following a single oral dose of PF-04457845 (0–40 mg). (**b**) Systems pharmacology model predictions of the effect on AEA levels following a single dose of 10 mg PF-04457845 assuming the absence (dotted line) or presence (solid line) of a FAAH-independent clearance process. The open circles show the mean experimental data from the healthy volunteer study (ref. 26). (**c**) Systems pharmacology model–predicted elevations of central nervous system CB1 receptor occupancy following a single dose of PF-04457845 (0.1–40 mg).

Systems pharmacology modeling to accelerate drug development and inform decision making for a sodium/glucose cotransporter 2 inhibitor in type 2 diabetes

Sodium/glucose cotransporter 2 (SGLT2) inhibitors are a class of glucose-lowering agents that provide a rational approach for the treatment of type 2 diabetes mellitus (T2DM);[29] SGLT2 is one of the more competitive therapeutic targets, with several inhibitors currently in development.[30,31] SGLT2 plays a major role in the reabsorption of glucose by the kidney,[32] and SGLT2 inhibition results in increased urinary glucose excretion (UGE) in both healthy subjects and subjects with T2DM,[33–36] leading to reductions in plasma glucose and glycosylated hemoglobin (HbA1c) levels with the additional benefit of a negative caloric balance, thereby combining glucose control with weight loss.[37,38] UGE provides the opportunity of a mechanistic biomarker for clinical assessment.

During the discovery and development of the SGLT2 inhibitor PF-04971729 (ertugliflozin), we continuously developed, used, and refined modeling and simulation tools to increase speed and enhance decision making in the program. These tools included biologically based PK–PD models in the nonclinical phase of the program,[39] model-based meta-analyses to assess comparative efficacy within[40] the class as well as against other antidiabetic agents,[41] and systems pharmacology models[42] to increase the confidence in the dose selection and trial outcome. This example will expand on use of the systems pharmacology model.

To inform the clinical development plan for ertugliflozin, we integrated the available data on the molecule, the physiologic understanding of the mechanism of action, and the published data on other SGLT2-targeting compounds in a comprehensive model. This allowed us to establish a link between the mechanistic biomarker UGE in healthy subjects and improvements in glycemic control, as measured by HbA1c, and body weight in longer-term studies in T2DM subjects. To facilitate this, we modeled SGLT2 inhibition using the Metabolism PhysioLab platform[43] (Entelos, Foster City, CA).

The Metabolism PhysioLab platform is a mathematical model of human T2DM pathophysiology consisting of several hundred ordinary differential and algebraic equations. The model is based on an extensive survey of published literature and represents the major physiological systems involved in the regulation of nutrient intake, utilization, storage, and disposal in health and disease. T2DM virtual patients are created by parametrically changing the underlying disease pathophysiology (e.g., insulin and glucose effects on metabolic pathways) to reflect the phenotypic diversity observed in T2DM clinical trials. The model has been validated by comparing virtual patient responses to published and proprietary data from a variety of clinical studies and trials, including oral glucose tolerance tests, clamp procedures, mixed-nutrient meal tests, nutrient and hormone infusions, and therapeutic interventions.[43]

A physiologically based representation of competitive SGLT2 inhibition was added to the model to account for the effect on glucose reabsorption in the proximal tubule. The model was updated using publicly disclosed information on the PK and UGE profile for experimental SGLT2 inhibitors in both healthy and T2DM subjects[33,34] from single- and multiple-dose studies with different meal protocols, and was validated by comparing predictions of efficacy against a published 12-week trial.[37] Phase I and phase II clinical trials were simulated (single- and multiple-ascending-dose trials in healthy and T2DM subjects, 12-week chronic dosing in T2DM subjects) and the results were compared with reported data for UGE, HbA1c, and body weight to calibrate the model. An E_{max} model was implemented to describe the renal glucose reabsorption rate in the presence of an SGLT2 inhibitor, and the

model parameters were calibrated to match the 24-h UGE data from single- and multiple-dose trials with other SGLT2 inhibitors. For 12-week phase II simulations, hypotheses related to disease progression and food intake were evaluated for consistency with published clinical HbA1c and body weight data. An increase in food intake in the presence of an SGLT2 inhibitor was required to match the weight loss reported in the 12-week clinical trial, representing a partial compensation for the negative energy balance resulting from the glycosuria.

Once representation of the comparator SGLT2 inhibitor was qualified, a simultaneous representation of SGLT2 inhibition with ertugliflozin was implemented in the platform. Food intake was implemented per clinical protocol. Modeled PK was introduced, and drug potency for ertugliflozin was adjusted in real time during the first-in-human study as PK and biomarker data with ertugliflozin became available to match UGE (**Figure 5a**). The resultant model was used to simulate the HbA1c dose–response relationship after 12 weeks of treatment with ertugliflozin to support dose selection for the phase II trial and the model-simulated response were subsequently corroborated by comparison to the observed clinical trial data (**Figure 5b**).

The modeling approach undertaken to support dose selection for the phase II study provided a quantitative link between UGE, the biomarker for the mechanism of action, and the long-term end points (HbA1c and body weight). Furthermore, this allowed successful prediction of changes in other circulating hormones (e.g., insulin levels) as well as providing grounds for simulating different dosing regimens and combination therapies. This information was used to successfully predict efficacy in T2DM patients from the observed first-in-human UGE data and to effectively project the phase II dose range with results from a single-dose escalation study in healthy subjects. This approach was one of the key drivers that allowed completion of the phase I and II clinical explorations in a shorter than expected time frame (14.6 months).

Adaptive dose-finding phase IIb study designed using clinical trial simulations

A challenge for dose-finding trials of an anticoagulant is minimization of the serious risks of underdosing (thrombosis) and overdosing (major bleeding, MB) while exploring a wide enough dose range of the study drug to select an optimal phase III dose. PD 0348292 is an oral factor Xa inhibitor that was under investigation for prevention of venous thromboembolism (VTE) following total hip and knee replacement surgery. A modeling and simulation approach was undertaken to design a phase II dose-finding trial in total knee replacement surgery that would minimize the risk to patients of excessive VTE or MB but have a high probability of identifying a single phase III dose that would be expected to provide the appropriate balance between efficacy and safety. The objectives of the phase II study were therefore to (i) estimate the dose of PD 0348292 that is equivalent to the standard-of-care comparator, enoxaparin, for prevention of VTE in total knee replacement and (ii) characterize the dose–response relationship of PD 0348292 for VTE and MB. Exposure–response modeling and clinical trial simulations were used to leverage prior knowledge and evaluate the ability of alternative study designs to meet these objectives.

A PK–PD model was developed to link response in an *in vitro* PD assay (inhibition of thrombin generation) to clinical outcomes for comparator anticoagulant compounds, and the model was used to predict the VTE and MB dose–response relationships (and associated uncertainty) for PD 0348292, based on its response in the *in vitro* assay.[44] A database of study-level VTE and MB outcomes in hip and knee replacement surgery from the literature was used to characterize incidence of VTE and MB as a function of dose for 21 anticoagulant compounds via a logistic regression model that specified the same shape of the dose–response relationship but different potencies across compounds.[45] Data for 10 comparator compounds and PD 0348292

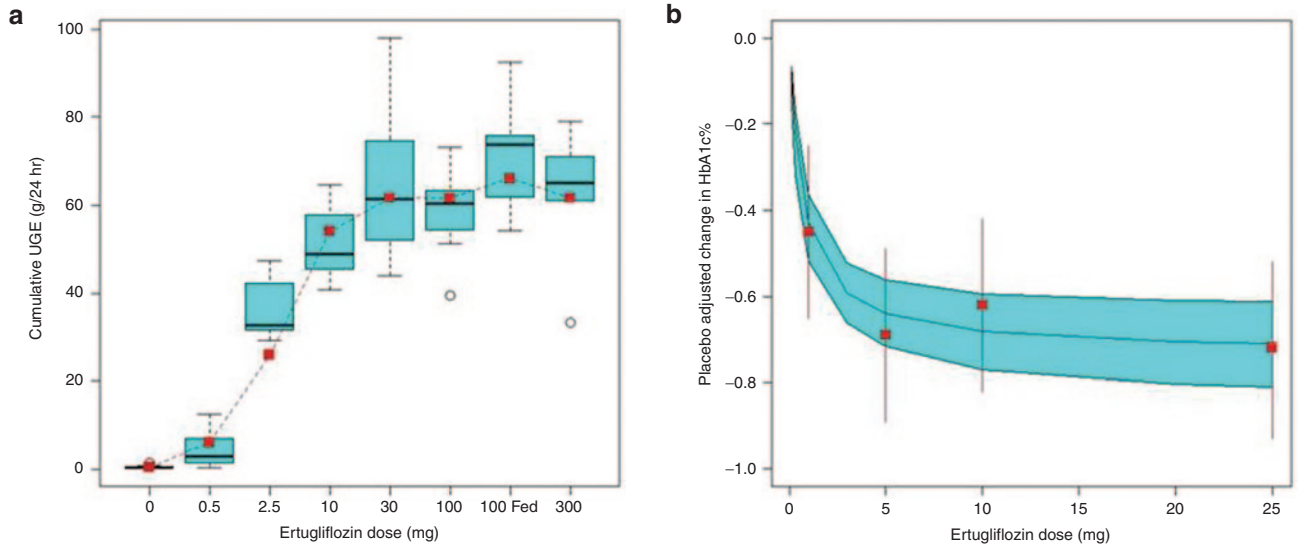

Figure 5 Systems pharmacology model prediction of the effect of ertugliflozin on (**a**) cumulative amount of urinary glucose excretion (UGE) in healthy subjects (red symbols, superimposed on observed data, represented by the box plots) and (**b**) prediction of the 12-week HbA1c results in type 2 diabetic subjects (shaded blue area is the 90% confidence interval (CI) of prediction, superimposed on observed data, red symbols, with observed 80% CI).

were generated in the PD assay, and clinical trial outcomes were available for 5 of these compounds.[44]

The VTE and MB dose–response relationships for PD 0348292 predicted by the models were quite uncertain (**Figure 6**) because there was no consistent relationship linking the inhibition of thrombin generation to clinical outcome across the compounds. The PD 0348292 dose equivalent to enoxaparin was estimated to be 1.1 mg (90% confidence interval: 0.076–16) for VTE and 1.5 mg (0.23–11) for MB. The models were used to simulate the outcome of each potential trial design 1,000 times.[44] Acceptable trial performance required that a single PD 0348292 dose be identified with both VTE and MB rates similar (ratio ≤ 1.3) to the comparator with a probability (power) ≥80% if there truly were a dose that satisfied those criteria. Factors that were considered in the evaluation of different study designs included sample size, analysis method (model-based dose–response analysis vs. pair-wise comparisons), inclusion of active control, randomization ratio between treatment arms, and adaptive dose modifications. For the adaptive designs, simulations were also used to evaluate various criteria for pruning and adding doses to ensure that (i) the overall power of the trial was maintained, (ii) the overall incidences of VTE and MB for PD 0348292 in the trial were similar to those for enoxaparin, and (iii) there was <5% cumulative chance that the dose equivalent to enoxaparin was pruned and ~70% cumulative chance that a nonequivalent dose was pruned.[44]

As a result of the clinical trial simulations, a randomized, active-controlled, parallel-group, adaptive dose-ranging study was conducted.[46] With the proposed study design, a fixed total sample size of 1,225 randomized subjects was estimated to provide at least 92% power that the selected dose of PD 0348292 would have true VTE and MB incidences within a factor of 1.3 of enoxaparin. A total of seven oral doses of PD 0348292 were studied (0.1–10 mg). The initial cohort of subjects was randomized to one of the lowest five doses of PD 0348292 or enoxaparin, in a 1:1:1:1:1:2 ratio. Over the course of the study, the data monitoring committee authorized discontinuation of the three lowest dose groups and addition of two higher dose groups on the basis of prespecified, periodic, model-based dose–response analyses of all available VTE and MB data.

A total of 1,411 subjects were randomized to treatment. The protocol was amended at the discretion of the data monitoring committee to increase the sample size by ~200 to allow for adequate sample size in the highest dose group because interim analyses were delayed slightly from the original plan. A statistically significant dose–response relationship ($P < 0.0001$) for incidence of VTE was observed. The relative risk for VTE was estimated with good precision (**Figure 6a**). On the basis of this dose–response model for efficacy, the dose of PD 0348292 equivalent to enoxaparin was estimated to be 1.16 mg q.d. (95% confidence interval: 0.56 mg–2.41 mg). The number of MB events was very low, and the dose–response relationship for MB was not statistically significant for PD 0348292 ($P = 0.6$, **Figure 6b**).

This study successfully used an adaptive design and a dose–response model-based analysis to safely assess PD 0348292 across a 100-fold dose range, whereas previous phase II studies of factor Xa inhibitors had evaluated a 4- to 12-fold dose range that was inadequate to discern a dose–response relationship for efficacy and/or resulted in doses being dropped for unacceptable bleeding.[47–50] On the basis of multiple prespecified dose decision analyses of the incidence of VTE and MB, doses associated with an unacceptable incidence of VTE were dropped early, ensuring minimum exposure of subjects to undesirable doses. Furthermore, doses considered to be safe based on predicted incidence of MB were added to the study, thereby allowing for safe exploration of the upper end of the dose range of PD 0348292. Characterization of the dose–response relationships for VTE and MB in this adaptive design study provided a robust scientific basis for phase III dose selection with ~2.5-fold fewer subjects than would be required in a conventional, parallel-arm, pairwise comparison study designed to achieve the same precision.

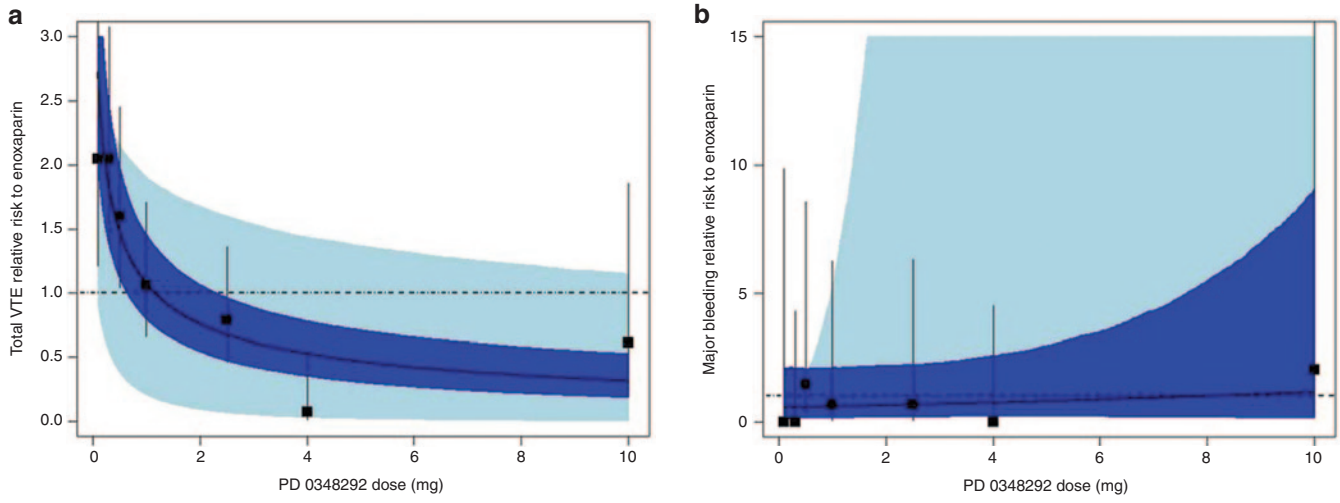

Figure 6 Observed relative risk of PD 0348292 vs. enoxaparin (symbols with 95% confidence intervals (CIs)) for (**a**) VTE and (**b**) MB and logistic regression model fit (solid line with dark blue area covering the 90% CI) in an adaptive phase II study. The light blue area covers the 90% CI before the trial based on the PK–PD model for inhibition of thrombin generation. MB, major bleeding; PK–PD, pharmacokinetics–pharmacodynamics; VTE, venous thromboembolism.

Application of MBDD in late clinical development: tofacitinib for rheumatoid arthritis

This example illustrates the prospective application of MBDD concepts to the late-stage development of tofacitinib, a novel, oral, small-molecule inhibitor of the Janus kinase (JAK) pathways, for the treatment of rheumatoid arthritis. Results from a POC trial demonstrated a high degree of efficacy but with side effects at all tested doses. The goal was to identify dose(s) for pivotal registration trials that would achieve a minimally acceptable product profile of similar efficacy as biologic injectables, with acceptable safety.

Tofacitinib is a JAK inhibitor. JAK enzymes transmit the signaling of several proinflammatory cytokines involved in the pathogenesis of rheumatoid arthritis through pairings of JAKs (e.g., JAK1/JAK3, JAK1/JAK2), and tofacitinib works by inhibiting the activities of these combinations, resulting in modulation of cellular processes of hematopoiesis and immune cell function.

The first evidence of efficacy in patients with rheumatoid arthritis was observed in a 6-week POC study of 5, 15, and 30 mg b.i.d. doses of tofacitinib and placebo.[51] All doses demonstrated efficacy as measured by the American College of Rheumatology (ACR) response criteria but were also associated with side effects such as dose-dependent changes in laboratory markers (e.g., decreased neutrophils). The challenge was to efficiently yet comprehensively characterize dose–response relationships to identify optimal dose(s) for confirmatory trials. This process began by gaining agreement with stakeholders on the key questions and setting quantitative and action-oriented objectives for the phase IIb program, as illustrated below.[3]

1. What do we need to know? Identify the lowest dose with at least 30% difference in ACR 20 response vs. placebo by week 12.
2. How sure do we want to be? We desire 80% probability that the true response for the model-estimated dose will be within ±20% of the target efficacy magnitude, i.e., 24–36%.
3. What are we willing to assume? A pharmacologically based, longitudinal E_{max} model will be applied; the dose range derived from the monotherapy POC study data will be applicable to combination treatment with methotrexate; prior distributions of parameters for the model-based analysis will be weakly informed by the POC study data.

Various longitudinal, dose–response models were developed, including an indirect latent variable response model, relating pharmacologically based models to categorical data.[52] The various models gave similar predictions of the data but showed differences when extrapolating to lower doses and later time points. Consequently, they were used as "data-generation" models to ensure that the design chosen had robust operating characteristics over a range of "true" relationships.[53] A similar approach was implemented to characterize decreases in absolute neutrophil counts. Because the neutropenia incidence data from the POC study were too sparse, modeling efforts were focused on characterizing neutrophil counts using indirect response and semimechanistic models[54] to provide a more stable basis for dose and time interpolation/extrapolation. Using clinical trial simulations, it was determined that the 10th percentile of the neutrophil count distribution was related to the risk of neutropenia and estimated with greater precision than the neutropenia incidence data, thereby providing an efficient way to eliminate doses with unacceptable neutropenia event rates predicted based on changes in continuous data.

Two 6-month, phase IIb studies were performed in which tofacitinib was administered either as monotherapy[55] or in combination with methotrexate.[56] Both studies evaluated placebo and tofacitinib doses of 1, 3, 5, 10, and 15 mg b.i.d. The sample sizes of these studies, totaling >800 patients, were larger than traditional phase II sample sizes because they were designed to support quantitative decision criteria aimed at identifying an optimal dose rather than statistical separation from placebo. Traditional pairwise comparisons between active doses would have necessitated a 70% increase in study size to achieve similar performance characteristics over the model-based approach.

Model-derived inferences, updated using Bayesian methods, were used to calculate the probability of technical success, i.e., the probability of achieving efficacy similar to that of standard of care.[57,58] As predicted from the POC study, changes in neutrophils and predicted incidence of neutropenia were within acceptable limits and therefore not considered to limit the dose range under consideration for phase III trials. However, dose-dependent changes in hemoglobin levels were noted. An empirical, longitudinal model was applied to capture the relationship between dose and hemoglobin levels. The probability that the incidence of clinically important anemia (defined as >2 g/dl decrease from baseline in hemoglobin or absolute value <8 g/dl) will not exceed 5% above placebo over 6 months of treatment was calculated.

As shown in **Figure 7**, modeling based on the methotrexate combination study predicted that doses from 5 through 10 mg b.i.d. inclusive would meet both the desired efficacy and safety criteria of having ~50% or greater probability of achieving efficacy similar to standard of care, with anemia rates <5% above placebo. By contrast, a 3 mg dose had a 10% chance of achieving the ACR 70 target as compared with a 40% chance for the 5 mg dose. The choice of 5 and 10 mg doses was independently

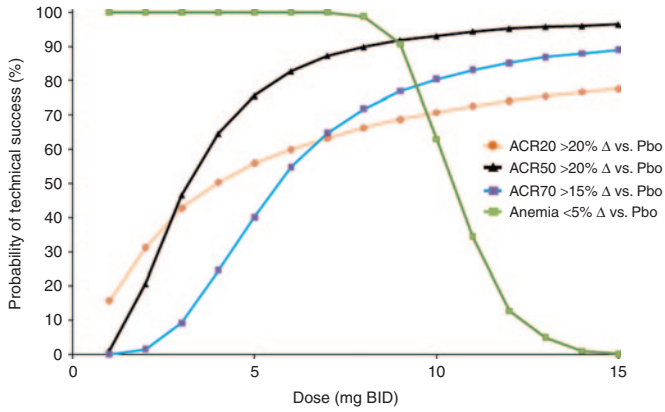

Figure 7 Tofacitinib—probability of achieving targeted differences vs. placebo. Solid symbols and lines represent model-based probability estimates for ACR responses and anemia. ACR, American College of Rheumatology; BID, twice daily; Pbo, placebo.

verified in the monotherapy phase IIb study. Doses ≥5 mg provided the requisite level of efficacy, whereas a 3 mg dose was considered clinically suboptimal, even though it separated from placebo.[55] Thus, the totality of the data justified the choice of 5 and 10 mg b.i.d. doses for phase III studies.

The results from the phase III program were consistent with these model predictions. The efficacy of 5 mg b.i.d. was as predicted (29% difference in ACR 20 rate vs. placebo across five phase III studies) and, more important, similar to that of standard of care (adalimumab).[59] The rates of anemia and neutropenia were low and considered manageable with appropriate clinical monitoring.

A prospective approach to (i) designing studies to stringent quantitative criteria, (ii) characterizing exposure–response relationships using well-established clinical outcome data in patient populations representative of the phase III program, and (iii) selecting doses based on efficacy and safety using probability of technical success as a common metric allowed demonstration of a positive benefit:risk profile with the desired product attributes. Tofacitinib of 5 mg b.i.d. was approved in 2012 by the US Food and Drug Administration for the treatment of moderately to severely active rheumatoid arthritis.

DISCUSSION

In **Figure 1**, we illustrated the five highly interdependent rationales (for selected pathway, target, molecule, dose regimen, and patients) that in combination derive an overall likelihood that a compound will become a commercially viable product. We believe that specific elements of MBDD can be applied in order to (i) determine a (numerical) level of confidence in each specific foundational rationale and (ii) potentially increase their magnitude. Historically, the clinical pharmacology and pharmacometrics community has exerted most influence through applying MBDD in the domain of benefit/risk, and indeed all the examples of our 2007 article addressed benefit/risk in one form or another. This will continue to be a core area for our community to influence because we have acquired experience and expertise and we have an evidence base of favorable impacts. In this article, example 5 follows a similar vein in that it reinforces the point that quantifying the probability of achieving the required product profile is a critical component of robust phase III strategy, study designs, and decision making. To this effect, we have emphasized the importance of comprehensive knowledge management strategies and have demonstrated that it is possible to reverse the trend of increasing phase III (and IV) study failure. Example 4 also addresses benefit/risk, although by different means. In this example the actual study design adopted was used as a "risk-mitigation strategy" due to its adaptive nature (because the accumulated evidence base gathered before designing the trial highlighted the high level of uncertainty in determining a favorable clinical dose range).

We have been unable to demonstrate an increase in phase II study success. To address this situation, we need to consider what factors are most impactful for phase II attrition, which, of course, may be different from the factors impacting phase III. By the time a compound is administered to humans, the benefit/risk attributes are largely predetermined (and therefore amenable to prediction) because we have already consciously selected both the particular target and compound to study. The role clinical development serves is to "simply" uncover the compounds' inherent benefit/risk properties (be they favorable or not) sooner or later within the clinic. Consequently, what would greatly increase the robustness of any prediction would be the depth of our understanding of the properties of the specific human target and how well preclinical compound-related attributes can inform the resultant human compound-related attributes. In the latter case (preclinical to clinical translation), the clinical pharmacology and pharmacometrics community, together with other groups, has acquired both experience and expertise, and we can again demonstrate an evidence base of favorable impacts. In the former case, this is an emerging area for our community because we see this as an important evolution of MBDD. We believe that if specific quantitative human target/system properties were to be more routinely integrated into our models, then our ability to predict human responses from data gathered in the preclinical space would be greatly enhanced. This is the rationale behind incorporating human systems biology and pharmacology domains into **Figure 1** because results in animal models of disease are often not predictive of efficacy in humans.[60] A more complete and comprehensive understanding of human biology and pathophysiology can be achieved via more mechanistic or system models.[61] By initiating such models early in the discovery stage, they can be propagated and updated through the continuous accumulation of experimental data, across systems, throughout discovery and development, thereby increasing confidence in the selected target.

The first three examples in this article illustrate the role that systems pharmacology models can now play. In example 1, we were able to determine the (non)viability of the target of interest without the need for any additional human studies. Example 2 also demonstrated the (non)viability of the target of interest, verified by a small amount of clinical data. What both these examples share is that they underwrote no go decisions for which we were able to be confident that we had thoroughly tested each target. Indeed, a methodical examination of the systems model described in example 2 afforded some insights on alternative targets/pathways that may be worthy of further investigations. The impact of example 3 was quite different. In this case, we were able to use our integrated models that captured pathology, physiology, and the competitive landscape to progress the compound to phase II extremely quickly (go with confidence).

In **Figure 1**, the first and last domains (pathway and effectiveness/reimbursement), although at different ends of the "spectrum," have a common attribute in that the volume of available data from each domain poses significant challenges for the types of methodological and analytical approaches commonly adopted by the clinical pharmacology and pharmacometrics community. In this respect the term "data-driven models" should not be regarded as a weakness in these data inferential capacities because the specific data types, such as cellular and subcellular functional responses, -omics, epidemiology, real-world data, and patient-reported outcomes, are abundant and valuable sources of relevant information. Both the systems biology and pharmacoeconomic

communities have adopted novel and interesting approaches to address their "data problems," and when the publications emanating from these two communities are examined, we (clinical pharmacology and pharmacometrics) can often see opportunities for complementary and mutually beneficial collaborations.[61,62] We strongly support collaborations between these different disciplines because they represent the next frontier for MBDD and an important opportunity to impact drug discovery, development, registration, and access to new medicines. Certainly, the pharmacoeconomic community regularly does something that clinical pharmacology and pharmacometrics do not, in that they publish sophisticated, integrated, relevant, and impactful data analyses in the journals of prescribers and payers.[63,64] Within the effectiveness and reimbursement arena, we consider these examples to characterize the current "best practice" application of MBDD. Because they emanate from the pharmacoeconomic community, this should serve to further emphasize our opportunity for important future collaboration.

The focus of this article was to highlight the nature of change within a pharmaceutical research and development organization in order to increase MBDD utilization. External influences can, and do, affect the rate of these internal changes. How worldwide regulatory agencies position themselves in terms of "supporting" MBDD can be important, and it is encouraging to witness recent developments within the European Medicines Agency[65] complementing those of the US Food and Drug Administration.[66]

Looking forward, all the aforementioned quantitative groups, including our own, are likely to face similar challenges in that it will not be the lack of opportunities, or suitable data, that will be our challenge but rather the lack of computational methods or modeling techniques or structures that can appropriately accommodate the wealth of available relevant data. Overemphasis on techniques and approaches to *collect* even more data, without a matching increase in our capacity and ability to *connect* the data (after A.F. Cohen, Leiden, The Netherlands) in an appropriate manner, will undoubtedly lead to missed opportunities. However, the continued evolution of MBDD in order to enable and enhance its relevance in this dynamic landscape will mean that MBDD will continue to offer a rational approach to efficiently accelerate drug development.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/cpt

ACKNOWLEDGMENTS

The authors acknowledge that the studies and impact presented in this article were the result of the efforts of hundreds of dedicated scientists and clinicians involved in the discovery and development of new medicines at Pfizer. In particular, the authors appreciate the key contributions of the following Pfizer colleagues: Mohan Beltangady, Jack Cook, Daniel Meyer, Alan Clucas, Charles Knirsch, Pankaj Gupta, Mathew Hutmacher, Kenneth Kowalski, Huaming Tan, Jonathan French, Neal Thomas, Meg Bennetts, Hannah Jones, David-Olivier Azulay, Avijit Ghosh, Nahor Haddish-Berhane, Cynthia Musante, Jeff Trimmer, Paul DaSilva-Jardine, Tristan Maurer, Neeta Amin, and William Denney. We also thank Oleg Demin, Evgeny Metelkin, and Natalia Bagrova, Institute for Systems Biology, Moscow, Russia; Peter Deuflhard, Susanna Roblitz, and Claudia Stotzel, Computational Systems Biology Group, Zuse Institute Berlin, Berlin, Germany; and Matt Riggs, Metrum Research Group, Tariffville, Connecticut.

CONFLICT OF INTEREST

All the authors are or were employees of Pfizer, with the exception of J.W.M., who is employed by Quantitative Solutions and has received consulting fees from Pfizer.

1. Lalonde, R.L. *et al*. Model-based drug development. *Clin. Pharmacol. Ther*. **82**, 21–32 (2007).
2. van der Graaf, P.H. CPT: pharmacometrics and systems pharmacology. *CPT: Pharmacomet. Syst. Pharmacol*. **1**, e8 (2012).
3. Sheiner, L.B. Learning versus confirming in clinical drug development. *Clin. Pharmacol. Ther*. **61**, 275–291 (1997).
4. Mahmood, T.A. & Templeton, A. Prevalence and genesis of endometriosis. *Hum. Reprod*. **6**, 544–549 (1991).
5. Surrey, E.S. & Hornstein, M.D. Prolonged GnRH agonist and add-back therapy for symptomatic endometriosis: long-term follow-up. *Obstet. Gynecol*. **99**, 709–719 (2002).
6. Pierce, S.J., Gazvani, M.R. & Farquharson, R.G. Long-term use of gonadotropin-releasing hormone analogs and hormone replacement therapy in the management of endometriosis: a randomized trial with a 6-year follow-up. *Fertil. Steril*. **74**, 964–968 (2000).
7. Hornstein, M.D., Yuzpe, A.A., Burry, K.A., Heinrichs, L.R., Buttram, V.L. Jr & Orwoll, E.S. Prospective randomized double-blind trial of 3 versus 6 months of nafarelin therapy for endometriosis associated pelvic pain. *Fertil. Steril*. **63**, 955–962 (1995).
8. Franke, H.R., van de Weijer, P.H., Pennings, T.M. & van der Mooren, M.J. Gonadotropin-releasing hormone agonist plus "add-back" hormone replacement therapy for treatment of endometriosis: a prospective, randomized, placebo-controlled, double-blind trial. *Fertil. Steril*. **74**, 534–539 (2000).
9. Fernandez, H., Lucas, C., Hédon, B., Meyer, J.L., Mayenga, J.M. & Roux, C. One year comparison between two add-back therapies in patients treated with a GnRH agonist for symptomatic endometriosis: a randomized double-blind trial. *Hum. Reprod*. **19**, 1465–1471 (2004).
10. Riggs, M.M., Bennetts, M., van der Graaf, P.H., Martin, S.W. Integrated pharmacometrics and systems pharmacology model-based analyses to guide GnRH receptor modulator development for management of endometriosis. *CPT: Pharmacomet. Syst. Pharmacol*. **1**, e11 (2012).
11. Nafarelin for endometriosis: a large-scale, danazol-controlled trial of efficacy and safety, with 1-year follow-up. The Nafarelin European Endometriosis Trial Group (NEET). *Fertil. Steril*. **57**, 514–522 (1992).
12. Adamson, G.D., Kwei, L. & Edgren, R.A. Pain of endometriosis: effects of nafarelin and danazol therapy. *Int. J. Fertil. Menopausal Stud*. **39**, 215–217 (1994).
13. Peterson, M.C. & Riggs, M.M. A physiologically based mathematical model of integrated calcium homeostasis and bone remodeling. *Bone* **46**, 49–63 (2010).
14. Röblitz, S. *et al*. A mathematical model of the human menstrual cycle for the administration of GnRH analogues. *J. Theor. Biol*. **321**, 8–27 (2013).
15. Di Marzo, V. Inhibitors of endocannabinoid breakdown for pain: not so FA(AH) cile, after all. *Pain* **153**, 1785–1786 (2012).
16. Sasso, O. *et al*. Peripheral FAAH inhibition causes profound antinociception and protects against indomethacin-induced gastric lesions. *Pharmacol. Res*. **65**, 553–563 (2012).
17. Roques, B.P., Fournié-Zaluski, M.C. & Wurm, M. Inhibiting the breakdown of endogenous opioids and cannabinoids to alleviate pain. *Nat. Rev. Drug Discov*. **11**, 292–310 (2012).
18. Johnson, D.S. *et al*. Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett*. **2**, 91–96 (2011).
19. Skaddan, M.B. *et al*. The synthesis and *in vivo* evaluation of [18F]PF-9811: a novel PET ligand for imaging brain fatty acid amide hydrolase (FAAH). *Nucl. Med. Biol*. **39**, 1058–1067 (2012).
20. Ahn, K. *et al*. Mechanistic and pharmacological characterization of PF-04457845: a highly potent and selective fatty acid amide hydrolase inhibitor that reduces inflammatory and noninflammatory pain. *J. Pharmacol. Exp. Ther*. **338**, 114–124 (2011).
21. Li, G.L. *et al*. Assessment of the pharmacology and tolerability of PF-04457845, an irreversible inhibitor of fatty acid amide hydrolase-1, in healthy subjects. *Br. J. Clin. Pharmacol*. **73**, 706–716 (2012).
22. Morgan, P. *et al*. Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discov. Today* **17**, 419–424 (2012).

23. Benson, N., van der Graaf, P.H. & Peletier, L.A. Cross-membrane signal transduction of receptor tyrosine kinases (RTKs): from systems biology to systems pharmacology. *J. Math. Biol*. **66**, 719–742 (2013).
24. Huggins, J.P., Smart, T.S., Langman, S., Taylor, L. & Young, T. An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of the knee. *Pain* **153**, 1837–1846 (2012).
25. Benson, N. *et al*. Systems pharmacology of the nerve growth factor pathway; use of a systems biology model for the identification of key drug targets using sensitivity analysis and the integration of physiology and pharmacology. *Interface Focus* **3**, 2–9 (2013).
26. Geerts, H., Spiros, A., Roberts, P. & Carr, R. Has the time come for predictive computer modeling in cns drug discovery and development? *CPT: Pharmacomet. Syst. Pharmacol*. **1**, e16 (2012).
27. Hendriks, B.S. *et al*. Multiscale kinetic modeling of liposomal doxorubicin delivery quantifies the role of tumor and drug-specific parameters in local delivery to tumors. *CPT: Pharmacomet. Syst. Pharmacol*. **1**, e15 (2012).
28. Lippert, J. *et al*. A mechanistic, model-based approach to safety assessment in clinical development. *CPT: Pharmacomet. Syst. Pharmacol*. **1**, e13 (2012).
29. Abdul-Ghani, M.A. & DeFronzo, R.A. Inhibition of renal glucose reabsorption: a novel strategy for achieving glucose control in type 2 diabetes mellitus. *Endocr. Pract*. **14**, 782–790 (2008).
30. Carpino, P.A. & Goodwin, B. Diabetes area participation analysis: a review of companies and targets described in the 2008 - 2010 patent literature. *Expert Opin. Ther. Pat*. **20**, 1627–1651 (2010).
31. Abdul-Ghani, M.A., Norton, L. & DeFronzo, R.A. Efficacy and safety of SGLT2 inhibitors in the treatment of type 2 diabetes mellitus. *Curr. Diab. Rep*. **12**, 230–238 (2012).
32. Kanai, Y., Lee, W.S., You, G., Brown, D. & Hediger, M.A. The human kidney low affinity Na+/glucose cotransporter SGLT2. Delineation of the major renal reabsorptive mechanism for D-glucose. *J. Clin. Invest*. **93**, 397–404 (1994).
33. Komoroski, B. *et al*. Dapagliflozin, a novel SGLT2 inhibitor, induces dose-dependent glucosuria in healthy subjects. *Clin. Pharmacol. Ther*. **85**, 520–526 (2009).
34. Komoroski, B., Vachharajani, N., Feng, Y., Li, L., Kornhauser, D. & Pfister, M. Dapagliflozin, a novel, selective SGLT2 inhibitor, improved glycemic control over 2 weeks in patients with type 2 diabetes mellitus. *Clin. Pharmacol. Ther*. **85**, 513–519 (2009).
35. Hussey, E.K. *et al*. Single-dose pharmacokinetics and pharmacodynamics of sergliflozin etabonate, a novel inhibitor of glucose reabsorption, in healthy volunteers and patients with type 2 diabetes mellitus. *J. Clin. Pharmacol*. **50**, 623–635 (2010).
36. Sha, S. *et al*. Canagliflozin, a novel inhibitor of sodium glucose co-transporter 2, dose dependently reduces calculated renal threshold for glucose excretion and increases urinary glucose excretion in healthy subjects. *Diabetes. Obes. Metab*. **13**, 669–672 (2011).
37. List, J.F., Woo, V., Morales, E., Tang, W. & Fiedorek, F.T. Sodium-glucose cotransport inhibition with dapagliflozin in type 2 diabetes. *Diabetes Care* **32**, 650–657 (2009).
38. Rosenstock, J. *et al*.; Canagliflozin DIA 2001 Study Group. Dose-ranging effects of canagliflozin, a sodium-glucose cotransporter 2 inhibitor, as add-on to metformin in subjects with type 2 diabetes. *Diabetes Care* **35**, 1232–1238 (2012).
39. Maurer, T.S. *et al*. Pharmacodynamic model of sodium-glucose transporter 2 (SGLT2) inhibition: implications for quantitative translational pharmacology. *AAPS J*. **13**, 576–584 (2011).
40. Denney, W.S. & Nucci, G. Model-based meta-analysis of HbA1c, weight, and FPG in type 2 diabetes: Focus on SGLT2 inhibitors. American Conference of Pharmacometrics, San Diego, CA, 3–6 April 2011.
41. Mandema, J., Sweeney, K., Terra, S. & Sahasrabudhe, V. Model-based meta-analysis of the HbA1c lowering effect of PF-04971729, a sodium glucose co-transporter-2 inhibitor (SGLT2i), in comparison with other SGLT2i and anti-diabetic agents (ADA). *Diabetes* **61** (suppl. 1), 1015 (2012).
42. Nucci, G. *et al*. Quantitative human pharmacology modeling to accelerate SGLT2i drug development. American Conference on Pharmacometrics, San Diego, CA, 3–6 April 2011.
43. Brazhnik, P., Hall, K., Polidori, D., Siler, S.Q. & Trimmer, J.K. Method and apparatus for computer modelling diabetes. US Patent Application 10/040,373 (2002).
44. Boyd, R. *et al*. Modeling and simulation facilitated design of an adaptive phase 2 dose-finding study for PD 0348292, a novel FXa inhibitor. *Clin. Pharmacol. Ther*. **83**, S61 (2008).
45. Mandema, J.W., Boyd, R.A. & DiCarlo, L.A. Therapeutic index of anticoagulants for prevention of venous thromboembolism following orthopedic surgery: a dose-response meta-analysis. *Clin. Pharmacol. Ther*. **90**, 820–827 (2011).
46. Cohen, A.T. *et al*. An Adaptive-Design Dose-Ranging Study of PD 0348292, a New Oral Factor Xa Inhibitor, for Thromboprophylaxis after Total Knee Replacement Surgery. *ASH Annual Meeting Abstracts* **112**, 980 (2008).
47. Turpie, A.G., Gallus, A.S. & Hoek, J.A.; Pentasaccharide Investigators. A synthetic pentasaccharide for the prevention of deep-vein thrombosis after total hip replacement. *N. Engl. J. Med*. **344**, 619–625 (2001).
48. Eriksson, B.I. *et al*.; ODIXa-HIP Study Investigators. Oral, direct Factor Xa inhibition with BAY 59-7939 for the prevention of venous thromboembolism after total hip replacement. *J. Thromb. Haemost*. **4**, 121–128 (2006).
49. Turpie, A.G. *et al*.; OdiXa-Knee Study Group. BAY 59-7939: an oral, direct factor Xa inhibitor for the prevention of venous thromboembolism in patients after total knee replacement. A phase II dose-ranging study. *J. Thromb. Haemost*. **3**, 2479–2486 (2005).
50. Lassen, M.R., Davidson, B.L., Gallus, A., Pineo, G., Ansell, J. & Deitchman, D. A phase II randomized, double blind, five-arm, parallel group, dose-response study of a new oral directly-acting factor Xa inhibitor, razaxaban, for the prevention of deep vein thrombosis in knee replacement surgery. *Blood* **102**, 41 (2003).
51. Kremer, J.M. *et al*. The safety and efficacy of a JAK inhibitor in patients with active rheumatoid arthritis: Results of a double-blind, placebo-controlled phase IIa trial of three dosage levels of CP-690,550 versus placebo. *Arthritis Rheum*. **60**, 1895–1905 (2009).
52. Hutmacher, M.M., Krishnaswami, S. & Kowalski, K.G. Exposure-response modeling using latent variables for the efficacy of a JAK3 inhibitor administered to rheumatoid arthritis patients. *J. Pharmacokinet. Pharmacodyn*. **35**, 139–157 (2008).
53. Krishnaswami, S. *et al*. Modeling and clinical trial simulation to design a dose-ranging study for CP-690,550 in rheumatoid arthritis patients. *Clin. Pharmacol. Ther*. **85** (suppl. 1), S61, PII–78 (2009).
54. Gupta, P., Friberg, L.E., Karlsson, M.O., Krishnaswami, S. & French, J. A semi-mechanistic model of CP-690,550-induced reduction in neutrophil counts in patients with rheumatoid arthritis. *J. Clin. Pharmacol*. **50**, 679–687 (2010).
55. Fleischmann, R. *et al*. Phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) or adalimumab monotherapy versus placebo in patients with active rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs. *Arthritis Rheum*. **64**, 617–629 (2012).
56. Kremer, J.M. *et al*. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. *Arthritis Rheum*. **64**, 970–981 (2012).
57. Tan, H., Gruben, D., French, J. & Thomas, N. A case study of model-based Bayesian dose response estimation. *Stat. Med*. **30**, 2622–2633 (2011).
58. Tofacitinib Arthritis Advisory Committee Meeting (FDA Advisory Committee, Washington, DC, 2012) < http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM304200.pdf>.
59. van Vollenhoven, R.F. *et al*.; ORAL Standard Investigators. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. *N. Engl. J. Med*. **367**, 508–519 (2012).
60. van der Worp, H.B. *et al*. Can animal models of disease reliably inform human studies? *PLoS Med*. **7**, e1000245 (2010).
61. Sorger, P.K. *et al*. *Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms* (National Institutes of Health, Bethesda, MD, 2011).
62. Lalonde, R.L. & Willke, R.J. Comparative efficacy and effectiveness: an opportunity for clinical pharmacology. *Clin. Pharmacol. Ther*. **90**, 761–763 (2011).
63. Pink, J., Lane, S., Pirmohamed, M. & Hughes, D.A. Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses. *BMJ* **343**, d6333 (2011).
64. Pink, J., Lane, S. & Hughes, D.A. Mechanism-based approach to the economic evaluation of pharmaceuticals: pharmacokinetic/pharmacodynamic/pharmacoeconomic analysis of rituximab for follicular lymphoma. *Pharmacoeconomics* **30**, 413–429 (2012).
65. Manolis, E., Rohou, S., Hemmings, R., Salmoson, T., Karlsson, M. & Milligan, P.A. The role of modelling and simulation in development and registration of medicinal products: output from the EFPIA/EMA modelling and simulation workshop. *CPT: Pharmacomet. Syst. Pharmacol*. **2**, e31(2013).
66. Lee, J.Y. *et al*. Impact of pharmacometric analyses on new drug approval and labelling decisions: a review of 198 submissions between 2000 and 2008. *Clin. Pharmacokinet*. **50**, 627–635 (2011).