

A latent variable approach for modeling categorical endpoints among patients with rheumatoid arthritis treated with golimumab plus methotrexate

Chuanpu Hu · Zhenhua Xu · Mahboob U. Rahman ·
Hugh M. Davis · Honghui Zhou

Received: 21 April 2010 / Accepted: 29 June 2010 / Published online: 16 July 2010
© Springer Science+Business Media, LLC 2010

Abstract The need for accurate exposure–response modeling is critical in the drug development process. Few methods are available for linking discrete endpoints, especially ordered categorical variables, to mechanistic (e.g., indirect response) models. Here we describe a latent-variable approach that is proposed in conjunction with an inhibitory indirect response model to link the placebo/comedication effect and drug exposure to the endpoints. The model is parsimonious, with desirable characteristics at initial timepoints, and allows simultaneous modeling of multiple endpoints that are categorically ordered. Application of the model is demonstrated with data from a phase 3 clinical trial of golimumab, a human IgG1 κ monoclonal antibody that binds with high affinity and specificity to tumor necrosis factor (TNF)- α , in patients with rheumatoid arthritis. The efficacy endpoints were 20, 50, and 70% improvement in the American College of Rheumatology criteria (ACR20, ACR50, and ACR70, respectively) as measures of improvement in disease severity. The modeling results were shown to be consistent by using either a sequential or simultaneous pharmacokinetic/pharmacodynamic modeling approach. The suitability of likelihood profiling and proper use of bootstrap methods in assessing parameter estimation precision are also presented. More accurate, parsimonious models with appropriately quantified uncertainty can facilitate better drug development decisions.

Keywords Discrete data · Population pharmacokinetic/pharmacodynamic modeling · NONMEM

C. Hu (✉) · Z. Xu · M. U. Rahman · H. M. Davis · H. Zhou
Pharmacokinetics, Modeling and Simulation, Biologics Clinical Pharmacology, Centocor Research and Development, Inc., 200 Great Valley Parkway, Malvern, PA 19355, USA
e-mail: chu25@its.jnj.com

M. U. Rahman
University of Pennsylvania School of Medicine, Philadelphia, PA, USA

Introduction

Exposure–response modeling facilitates optimal dosing selection, and its use is increasing in drug development. A widely used type of exposure–response model is that of indirect response models [1], which are usually physiologically plausible and thus lend confidence to the resulting predictions. However, they are designed to model continuous physiological endpoints, and their extension to discrete endpoints is not immediately clear. Hutmacher et al. [2] used mixed effect logistic regression to link indirect response model predictions to the binary endpoint of a 20% improvement in the American College of Rheumatology criteria (ACR20) [3] in patients with rheumatoid arthritis (RA). Recently, Lacroix et al. [4] provided a Markov transition approach for an exposure–response model with ACR20 as the pharmacodynamic variable. To our knowledge, these are the only approaches proposed for exposure–response modeling using the ACR20 response.

Logistic regression, which uses the logit transformation to link the predictors to the probability of the endpoint of interest, is the standard statistical tool to model categorical endpoints. The typical model takes the following form, used by both Hutmacher et al. [2] and Lacroix et al. [4]:

$$\text{logit}(\text{prob}[z = 1]) = \gamma + f_p(t) + f_d(t) + \eta \quad (1)$$

where $\text{prob}[z = 1]$ is the probability of the event of interest, γ relates to the baseline probability, $f_p(t)$ represents the time effect, $f_d(t)$ represents the drug or concentration effect, and η represents inter-subject variability. It can be convenient to interpret the sum of the terms $\gamma + f_p(t) + \eta$ as representing the effect in the placebo group.

In the model proposed by Hutmacher et al. [2], $\text{prob}[z = 1]$ is the probability of achieving ACR20, $f_p(t)$ has an exponential form that increases with time, and $f_d(t)$ is essentially the negative of the output of the indirect model. While the approach described their data well, the model predicts a nonzero probability of achieving an ACR20 response at $t = 0$, which may be viewed as theoretically undesirable. The Markov transition approach [4] focused on modeling the individual time course of the responder/nonresponder status. They also modeled the dropout probability, an aspect not investigated here, and thus we will disregard it for the ease of discussion. Their approach is equivalent to formulating $\text{prob}[z = 1]$ as the probability of achieving ACR20; however, at the subject level this is allowed to vary, depending on whether the subject is a responder or nonresponder. This applies to all terms in Eq. 1. There are therefore many more flexible components in their approach; for example, in their final model, $f_p(t)$ took the E_{\max} model form for responders, but was linear for nonresponders. The drug effect term $f_d(t)$ took the E_{\max} model form for both responders and nonresponders; however, the parameters were allowed to vary.

Here we propose an approach that can be more parsimonious than the one described by Hutmacher et al. [2] and has the desirable property of predicting $\text{prob}[z = 1] = 0$ when $t = 0$. We shall argue that it is more easily interpretable and may be better extrapolated than the Markov approach. The approach is demonstrated with data from a clinical phase III study of golimumab [5], a new human immunoglobulin G1 kappa (IgG1 κ) monoclonal antibody that binds with high affinity to TNF- α , was recently approved for the treatment of RA, psoriatic

arthritis, and ankylosing spondylitis [6], and has demonstrated linear pharmacokinetics [7]. Results from the likelihood profile and the bootstrap methods of assessing parameter estimation uncertainty are also compared with the NONMEM standard error of parameter estimates.

Methods

Data and information used for exposure–response modeling

Model development was carried out using data from a phase III placebo-controlled trial, GO-FORWARD, the design of which has been previously described [5]. It was a parallel group study in patients with active RA despite prior use of methotrexate (MTX) therapy. This analysis used data from a total of 302 patients who received MTX and either golimumab 50 mg, golimumab 100 mg, or placebo every 4 weeks through week 48. At week 16, patients who received placebo + MTX or golimumab 50 mg + MTX and had less than 20% improvement in swollen and tender joints had their dose adjusted to golimumab 50 mg + MTX or golimumab 100 mg + MTX, respectively. No dose adjustment was performed for patients who were originally assigned to receive golimumab 100 mg + MTX. At week 24, all patients in the placebo + MTX group crossed over to golimumab 50 mg + MTX. At least 10 serum golimumab concentration measurements per patient were scheduled to be taken over the course of the study.

Clinical response endpoint

Patients were assessed on whether ACR20, ACR50, or ACR70 responses were achieved at weeks 4, 8, 12, 14, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52. There were 2,340 ACR response data points from 171 patients in the combined golimumab + MTX group. The placebo + MTX group had 830 ACR response data points from 131 patients; among these, 42 had their dose adjusted at week 16 (early escape), and the ACR responses after week 20 were excluded for ease of graphing and interpretation. Among the ACR responses available for modeling, there were seven instances where all ACR20, ACR50, or ACR70 data were not collected and these were removed from the dataset, resulting in a total of 3,163 response points. The distribution of the time and response points in the treatment groups were as follows. In the placebo + MTX group, the number of response points were 131, 130, 130, 128, 128, 88, 86, 4, 2, and 1 for weeks 4, 8, 12, 14, 16, 20, 24, 28, 32, and 36, respectively. The decrease in number of responses at weeks 20 and 24 and the few present at weeks 28, 32, and 36 appeared to be due to clinical trial irregularities. In the golimumab + MTX groups, for all weeks between 4 and 52, the number of response points ranged between 84 and 86 in the golimumab 50 mg group and 75–85 in the golimumab 100 mg group.

A complexity arises, at least in theory, with modeling responses when treatment can be determined thereby, thus in this situation making the dose or treatment

random variables that covary with response. This issue is theoretically possible to address but practically complex. It has been studied to some extent by Beal [8], which was perhaps under-appreciated due to its complexity, and that no apparent serious biases of ignoring this complexity were suggested. In this manuscript, the convenient choice of treating all doses as given was used.

Pharmacokinetic model

A population pharmacokinetic (PK) model was developed using data from the combined golimumab + MTX group. Parameter estimation was implemented in the software NONMEM (Version 6, level 1, ICON Development Solutions, Elliott City, MD). Details of the PK study design and analysis are described elsewhere [9]. The model described the data adequately, and the result was consistent with a previous analysis using the confirmatory approach [10]. Empirical Bayesian parameter estimates were then used in the sequential PK/pharmacodynamic (PD) modeling approach using ACR response as discussed below. A detailed discussion on the sequential and simultaneous PK/PD approaches may be found in Zhang et al. [11].

Latent variable model for ACR response

The proposed approach assumes the existence of a latent variable, referred to as ACRL, that determines the disease condition, with smaller values indicating symptom improvement. The anti-TNF- α mechanism of action of golimumab suggests that the latent variable ACRL could follow a type I indirect response PK/PD model [1].

The model is provided as follows:

$$\frac{d \text{ACRL}(t)}{dt} = k_{\text{in}} \times IH - k_{\text{out}} \times \text{ACRL}(t) \quad (2)$$

where k_{in} is the formation rate of RA, and k_{out} is the amelioration remission rate. The combined placebo/MTX effect and the drug effect are assumed to contribute to the inhibition effect, IH , as follows:

$$IH = 1 - plb - drug, \quad (3)$$

The contribution of placebo/MTX effect, plb , was modeled empirically as follows:

$$plb = plb_{\text{max}} (1 - \exp(-k_{\text{plb}} \times t)) \quad (4)$$

where plb_{max} is the maximum placebo effect; t is the time after the first dose; and k_{plb} is the rate constant.

The contribution of drug effect to the inhibition of the formation rate, $drug$, was modeled with the E_{max} function:

$$drug = \frac{E_{\text{max}} \times C_p}{EC_{50} + C_p} \quad (5)$$

where E_{\max} is the maximum drug effect; EC_{50} is the concentration causing 50% of the maximum effect; and C_p is the concentration. In practice, it may be difficult to identify EC_{50} when exposure is much higher than the EC_{50} . In this situation it is helpful to consider the following reduced model, which could be viewed mathematically as taking the limit of $EC_{50} \rightarrow 0$:

$$\text{drug} = \begin{cases} 0, \text{ placebo} \\ E_{\max}, \text{ active} \end{cases} \tag{5'}$$

The usual physiological constraint requires that $0 \leq IH \leq 1$ in Eqs. 1 and 2 above. This can be achieved by perhaps the most common approach to constraining two positive variables that sum up to less than 1, by the reparameterization of parameters p_1 and p_2 as $p_1/(1 + p_1 + p_2)$ and $p_2/(1 + p_1 + p_2)$. We have proposed elsewhere an approach of relaxing the constraint to only $IH \leq 1$ [9]. In practice, data may not allow all parameters to be estimated; in which case it may be reasonable to assume that full inhibition is achievable, i.e., $plb_{\max} + E_{\max} = 1$.

It was further assumed that ACRr, where $r = 20, 50, \text{ or } 70$, is achieved whenever $r\%$ improvement from baseline has been achieved, i.e.,

$$[\text{ACRL}(\text{baseline}) - \text{ACRL}]/\text{ACRL}(\text{baseline}) \geq r$$

The latent variable ACRL is only defined up to a proportional constant. Therefore it was further assumed that $\text{ACRL}(\text{baseline}) = 1$. This leads to $k_{\text{in}} = k_{\text{out}}$ in Eq. 2. Thus, ACRr, where $r = 20, 50, \text{ or } 70$, is achieved whenever $\text{ACRL} \leq 1 - r/100$. Conceptually, simultaneous modeling of ACR20, ACR50, and ACR70 endpoints leads to a more effective use of data and thus more accurate and precise results.

Exposure–response model

We propose that the probability of achieving ACRr, with $r = 20, 50, \text{ or } 70$, can be modeled as

$$\text{logit}(\text{prob}[\text{ACRr} = 1]) = \text{logit}(1 - r/100) + \text{logit}(1 - \text{ACRL}) + \eta \tag{6}$$

where $\text{logit}(x) = \log(x/(1 - x))$ is the commonly used logit function. Mathematically, the term $\text{logit}(1 - \text{ACRL})$ may be viewed as the latent variable instead of ACRL, if preferred. This model may be motivated as follows. If r is allowed to vary on a continuous scale of 0–100 instead of 20, 50, or 70, the ease of achieving $r\%$ of ACR improvement is controlled by two factors: easing the criterion (reducing r , or equivalently, $r/100$) and reducing disease (ACRL). Since 100% improvement should be unlikely to achieve and 0% likely, it is desirable that $\text{prob}[\text{ACRr} = 1] = 0$ if $r = 100$ and 1 if $r = 0$, which is achieved by the term $\text{logit}(1 - r/100)$ in Eq. 6. The term $\text{logit}(1 - \text{ACRL})$ served a similar purpose.

Compared with the common logistic regression approach as in Eq. 1, this model has two advantages. First, at time 0, $\text{ACRL} = 1$ and therefore $\text{logit}(1 - \text{ACRL}) = -\infty$, making $\text{prob}[\text{ACRr} = 1] = 0$. This is desirable because at time 0, by definition, the ACRr response should not be achieved. Secondly, this model does not need to estimate any intercept parameters; the term $\text{logit}(1 - r/100)$ plays the role of intercept(s) in a natural sense and can be used simultaneously for ACR20, ACR50,

and ACR70, making the model more parsimonious. This may be desirable because categorical data are not as informative as continuous data and therefore may not allow many parameters to be reliably estimated. The term $\text{logit}(1 - r/100)$ could still be replaced by separate intercepts if needed.

Next, simultaneous modeling of ACR20, ACR50, and ACR70 responses was conducted as follows. Four outcomes are possible, and the associated probabilities are given below:

$$(1) \text{ achieving ACR70 : } \text{prob}(\text{ACR70} = 1) \quad (7)$$

$$(2) \text{ achieving ACR50 but not ACR70 : } \text{prob}(\text{ACR50} = 1 \text{ and } \text{ACR70} = 0) \\ = \text{prob}(\text{ACR50} = 1) - \text{prob}(\text{ACR70} = 1) \quad (8)$$

$$(3) \text{ achieving ACR20 but not ACR50 : } \text{prob}(\text{ACR20} = 1 \text{ and } \text{ACR50} = 0) \\ = \text{prob}(\text{ACR20} = 1) - \text{prob}(\text{ACR50} = 1) \quad (9)$$

$$(4) \text{ not achieving ACR20 : } \text{prob}(\text{ACR20} < 1) = 1 - \text{prob}(\text{ACR20} = 1) \quad (10)$$

All needed probabilities are given by Eq. 6.

Model implementation

The proposed models, i.e., Eq. 6 for ACR20, 50, or 70, or Eqs. 7–10 for joint modeling of ACR20, 50, and 70, were implemented in the same way as the standard approach of Eq. 1. The sequential PK/PD modeling approach was the primary method, during which the empirical Bayesian estimates for individual PK parameters were first obtained using observed golimumab concentrations and then fixed for the next step of estimating PD model parameters. The simultaneous PK/PD modeling approach, estimating all population PK/PD model parameters simultaneously from all observed golimumab concentrations and ACR measurements, was also used and results were compared with those from the sequential approach. All parameter estimations were implemented in NONMEM, using the LAPLACE method.

The likelihood profile method and bootstrap were used to provide additional assessments of estimation precision. With the likelihood profile method, the NONMEM objective function values (OFVs), which are approximately -2 times loglikelihood, were plotted against fixed parameter values. For appropriate interpretation of bootstrap results, it is important to choose stratification variables corresponding to the study randomization, which in the current case is the treatment group. Therefore the bootstrap samples were generated with the number of patients in each treatment group fixed as those in the original data. In order to reach reasonable precision, 2,000 runs were conducted. Boxplots of the distributions of the bootstrap estimates were generated. Visual predictive check (VPC) of the final model was also conducted by simulating 500 replicates of the dataset and comparing simulated and model-predicted ACR response frequencies over the treatment period.

Bootstrap use

We have stated elsewhere [9] the view that, despite the claim of “model validation” used at times, the ordinary use of the bootstrap method provides no benefit other than an alternative assessment of estimation precision. In particular, comparing mean or median bootstrap results to original model parameter estimates does not provide a valid estimate of bias. A concise mathematical proof is given below. Let D be the original dataset, and $F(D)$ be an unbiased estimator of parameter θ . Then, for the arbitrary value of B , the estimator $G(D) = [F(D) + B]$ has bias B . Also assume that N bootstrap datasets (D_i , $i = 1, \dots, N$) are generated from D , and the naïve procedure, intended for assessing the estimation bias of the estimator G , of comparing the mean bootstrap estimates and the original is denoted as NaiveBoot(G). Then

$$\begin{aligned} \text{NaiveBoot}(G) &= [\sum_i G(D_i)/N] - G(D) = \sum_i [F(D_i) + B]/N - [F(D) + B] \\ &= (\sum_i F(\text{data}_i)/N - F(D)) = \text{NaiveBoot}(F) \end{aligned}$$

Thus, evaluated by NaiveBoot(), the performance of the estimator G is identical to the unbiased estimator F . That is, NaiveBoot() has no ability to detect any bias, even if the bias is arbitrarily large. A similar argument holds if the median is used instead of the mean. This may explain why the practice of comparing the mean or the median bootstrap estimates and the original has never, to our knowledge, been reported in the literature to have found any evidence of bias.

Results

The sequential and simultaneous PK/PD modeling approach gave nearly identical PD parameter estimates, which was closer than expected from our previous experience. However, the simultaneous model estimation did not terminate successfully. Details of the results from the sequential approach are presented below.

For the PD modeling, the maximum placebo and drug effects could not be separately identified, thus the assumption of maximum inhibition, i.e., $\text{plb}_{\max} + E_{\max} = 1$, was made. In addition, EC_{50} could not be reliably estimated; the NONMEM standard error of the parameter estimate exceeded 100%, and the corresponding likelihood profile (not shown) was nearly flat, with a NONMEM OFV increase of less than 2 points and ranging from 0.1- to 4-fold of the point estimate. This may be attributed to an apparent plateau response for the two golimumab + MTX treatment groups in the study. Consequently, a reduced model (Eq. 5') was used instead of the concentration–effect model (Eq. 5), and the NONMEM OFV remained virtually unchanged. Model parameter estimates and their corresponding NONMEM standard errors are given in Table 1. Average observed ACR response frequencies and the predicted probabilities for the average patient at planned observation times, grouped by treatment, are shown in Fig. 1. The model appeared to adequately describe the average time trends of active treatments.

Table 1 Parameter estimates of the population pharmacokinetic/pharmacodynamic model

Parameters	Definition	Estimate	Standard error (CV%)
k_{out} (1/day)	Amelioration remission rate	0.0139	0.00389 (28.0)
E_{max}	Maximum drug effect	0.211	0.0558 (26.4)
k_{plb} (1/day)	Rate constant to reach the plateau	0.00133	0.000216 (16.2)
Var (η)	Variance of inter-subject variability on the logit scale	3.51	0.406 (11.6)

CV%, percent coefficient of variation

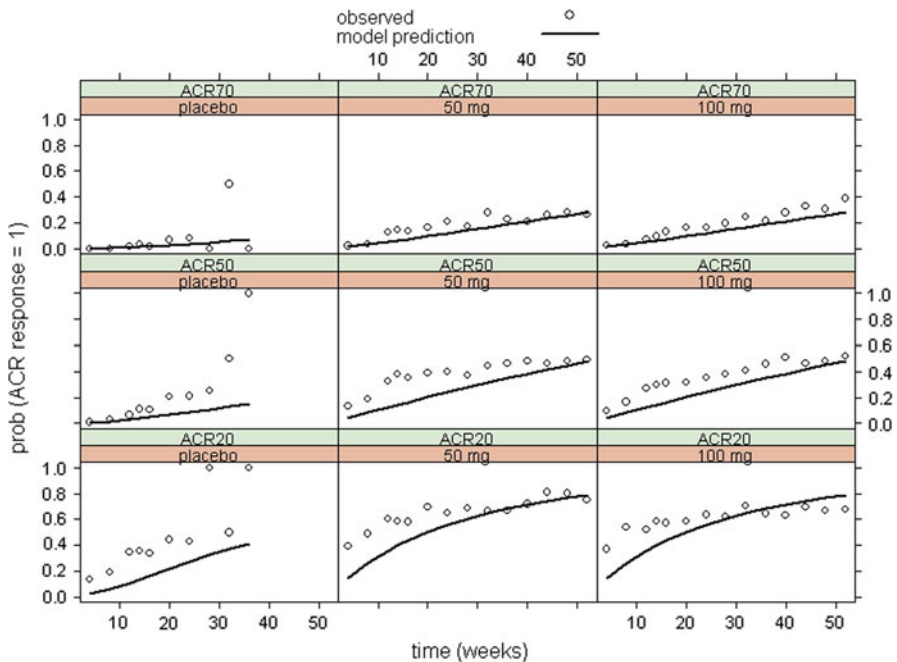


Fig. 1 Mean ACR20, ACR50, and ACR70 response frequencies at planned observation times versus the model predicted probabilities for the average patient, by treatment group. ACR20/50/70, 20/50/70% improvement in the American College of Rheumatology criteria

Estimation appeared reasonable, with most standard errors within the expected range of this type of data. It might appear that ACR70 had a delayed response relative to ACR20; however, as both endpoints measure the same mechanism, a delay would not be mechanistically justifiable. If such a delay exists, it would more likely be due to some type of inhomogeneity of the non-intercept terms in Eq. 6 in influencing ACR70 and ACR20 as ordered categories. The first step to investigate this would be to fit separate intercepts for ACR20, ACR50, and ACR70 in place of the term $\text{logit}(1 - r)$ in Eq. 6 as described below.

VPC results of the final model are shown in Fig. 2. The medians and 90% prediction regions of the simulated ACR response frequencies at planned

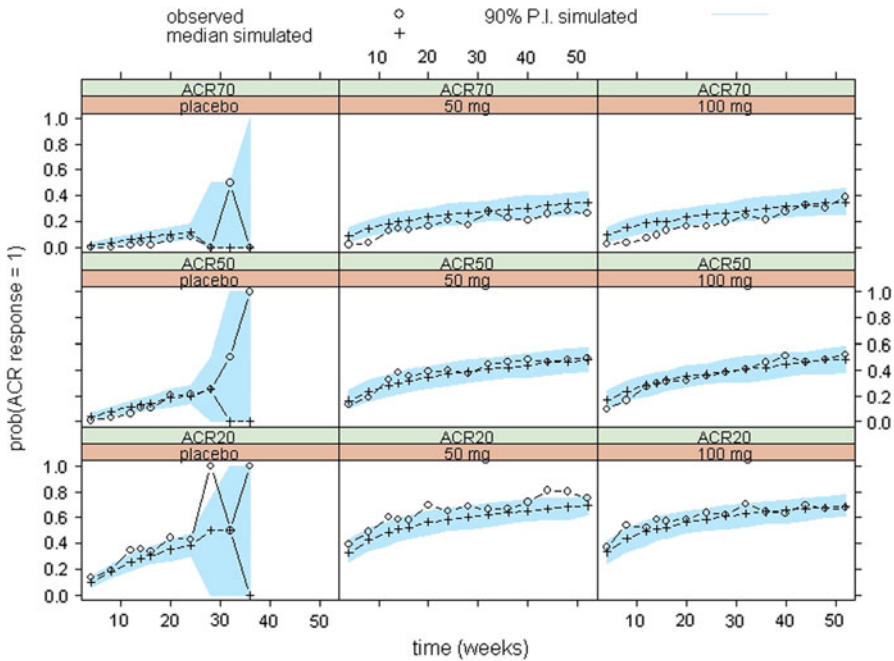


Fig. 2 Median model predictions at planned observation times and 90% prediction intervals (P.I.), in overlay with observed American College of Rheumatology (ACR) response frequencies. ACR20/50/70, 20/50/70% improvement in the American College of Rheumatology criteria

observation times are shown in overlay with the observed ACR response frequencies. The observed frequencies generally fell within the prediction intervals, although the model appeared to slightly underpredict ACR20 and overpredict ACR70. This might suggest that the model may be improved by fitting separate intercepts. To investigate the potential misfits suggested by Figs. 1 and 2 above, a model fitting separate intercepts for ACR20, ACR50, and ACR70 in place of the term $\text{logit}(1 - r)$ in Eq. 6 was also investigated. Precision of intercept estimates was poor, with a coefficient of variation greater than 50%. In addition, VPC results were also poor (Fig. 3), with predicted time trends considered clinically unrealistic. This indicated that the model was over-parameterized.

Comparison of Figs. 1 and 2 shows some discrepancies between the two types of model predictions, i.e., the model predicted probabilities for the average patient and the median of the VPC results. This, in essence, is due to the difference between the median and mean of the probability distribution represented by the model; the distribution is more skewed when the median, i.e., the predicted probability of the average subject, is close to 0 or 1. Overall, the final model could be viewed as a reasonable description of the data.

For ease of comparing different methods of estimating variability, all parameters were displayed on a proportional scale by normalizing them to (i.e., dividing by) the original model parameter estimate values given in Table 1. The likelihood profile results are shown in Fig. 4. It appears that, with an approximate 90% confidence,

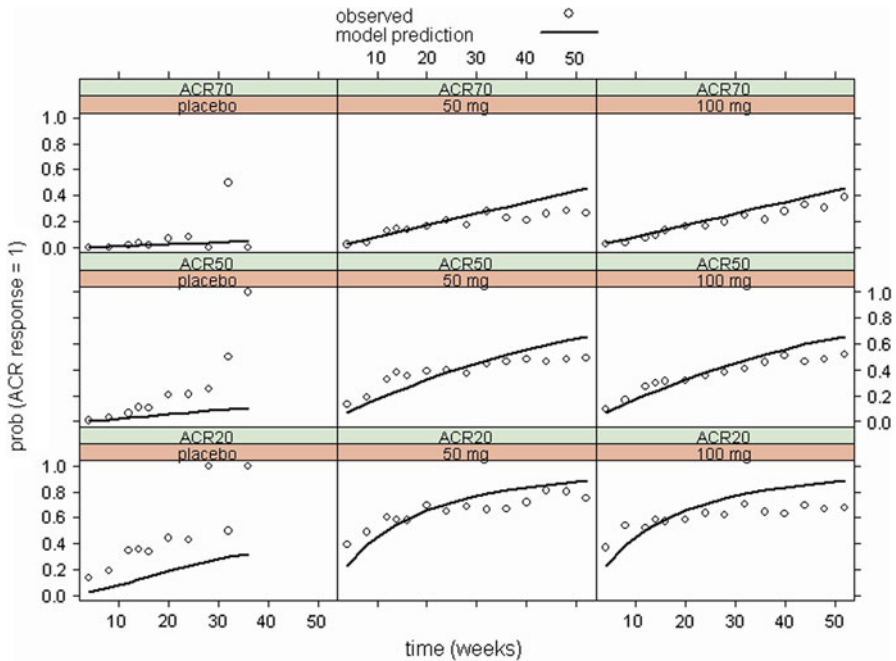


Fig. 3 Over-parameterized model fitting three separate intercepts: mean ACR20, ACR50, and ACR70 response frequencies at planned observation times versus those of model predictions by treatment group. ACR20/50/70, 20/50/70% improvement in the American College of Rheumatology criteria

k_{plb} could be determined within 50%, and E_{max} and k_{out} could be determined within 100%. The bootstrap results are given in Fig. 5. The edges of k_{plb} are within 50%, and the edges of E_{max} and k_{out} are within 100%. Comparison of the PK NONMEM parameter estimates (Table 1), the likelihood profile plots (Fig. 4), and the bootstrap analysis (Fig. 5) shows that similar conclusions can be drawn from the three methods of assessing parameter estimation uncertainty.

Discussion

A latent variable approach is proposed here along with an inhibitory indirect response model to describe the time course of ACR responses. This approach has some advantages over the existing models in patients with RA. Currently, the standard method is the mixed effect logistic regression [2] given in Eq. 1. Because at baseline (time = 0) both the placebo and drug effect are equal to 0, the fixed intercept term γ dictates a non-zero $\text{prob}[z = 1]$ which is the probability of achieving ACR20, which is not theoretically possible. In contrast, with our approach given in Eq. 6, at time 0 the term $\text{logit}(1 - \text{ACRL}) = -\infty$, leading to $\text{prob}[\text{ACR} = 1] = 0$. In addition, modeling multiple endpoints (ACR20, ACR50, ACR70) simultaneously has advantages over modeling a single one (e.g., ACR20) alone because more information is used. In this situation, adopting the approach of

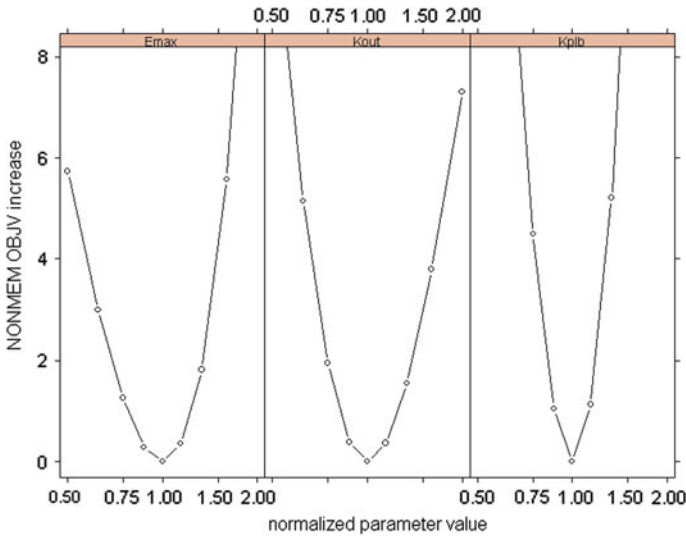


Fig. 4 Likelihood profiles for the fixed-effect parameters varying from 0.5 to 2 times of their estimates. The maximum of the plot range of OFV increase was truncated at 7.88 for ease of visibility. The value 1 indicates the original parameter estimate(s), for which the change of OFV is zero by definition. As normalized parameter values differ from 1, steep increases in the corresponding OFVs for fixed parameters indicate good estimation precision. *OFV* objective function value

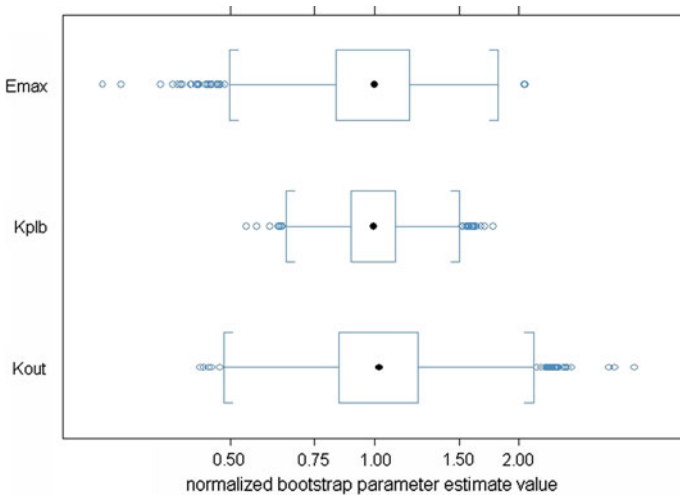


Fig. 5 Box-whisker plots of bootstrap estimates normalized by their original model estimates. Medians are in the center, black boxes contain 50% data, and the outside bounds are 1.5 times the inter-quartile ranges. Data beyond outside edges were interpreted as potential outliers

Eq. 1 would require multiple intercept terms to be used, one for each category. In contrast, the term $\text{logit}(1 - r/100)$ plays the role of intercepts but does not require any parameters to be estimated. This can result in efficient estimation and may be especially beneficial when information is sparse.

A more recently proposed approach uses the Markov transition [4]. We shall restrict the discussion to the situation when dropouts are few and thus the influence may be neglected. In this case, the Markov approach essentially still uses the form of Eq. 1 but allows all effect terms and parameters to change, depending on whether the patient is currently a responder. By describing the individual time course as transitions between the responder/nonresponder states, it has the appeal of accounting for within-subject correlation of the observations. However, complexities also arise. First, the model formulation implicitly assumes a time homogeneity in that the time effect on the transition probabilities appeared only indirectly through drug concentration terms. This may not be realistic as the model predicts nonzero probabilities of transition to a different state even as the time interval length approaches 0, in which case one would not expect any chance of change in the responder status. This implies that Markov models may have difficulties extrapolating to situations with different dosing and observation schedules. Secondly, because in principle the responder status is already related to the placebo and drug exposure, formulating the model, i.e., placebo and drug effect terms, to be conditional on the responder status may, at least in theory, create confounding complexities in interpreting the model terms. Finally, splitting the model terms, e.g., placebo or drug effect, to depend on current responder/nonresponder status makes it difficult to understand how each model term influences the overall response rate. This may create interpretation difficulties of the model, in light of Simpson's paradox (see e.g., Freedman, et al. [12]) which shows that a trend present in different groups can be reversed when the groups are combined. In contrast, the mixed effect logistic regression approach used here predicts, at any time point, directly the probability of a positive response of the average subject. Its relationship with the frequency of positive response of the population, which is the clinical trial endpoint, can be more easily discerned.

In the GO-FORWARD study, exposures in the active golimumab dose groups appeared to have reached a plateau in the exposure–response curve, making the relationship difficult to identify. Proper identification of exposure–response modeling allows prediction of trial outcomes under alternative dosing regimens, and the possibility should be explored. The proposed approach provides a means for this exploration.

Mechanistically interpretable approaches have more advantages over empirical approaches, such as cumulative AUC, as elegantly stated by Hutmacher et al. [2]. Our approach can be easily extended to model any other type of ordered categorical endpoints and to any other type of PK/PD model such as the other indirect response model types.

Finally, proper use of model evaluation methods, particularly bootstrap, can be important for appropriate results interpretation.

Acknowledgments The authors thank Rebecca E. Clemente, PhD, and Robert Achenbach of Centocor Ortho Biotech Services, LLC for their excellent editorial support. This study was funded by Centocor Research and Development, Inc.

References

1. Krzyzanski W, Jusko WJ (1998) Integrated functions for four basic models of indirect pharmacodynamic response. *J Pharm Sci* 87:67–72
2. Hutmacher MM, Krishnaswami S, Kowalski KG (2008) Exposure-response modeling using latent variables for the efficacy of a JAK3 inhibitor administered to rheumatoid arthritis patients. *J Pharmacokinet Pharmacodyn* 35:139–157
3. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, Katz LM, Lightfoot R Jr, Paulus H, Strand V, Tugwell P, Weinblatt M, Williams HJ, Wolfe F, Kieszak S (1995) American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 38:727–735
4. Lacroix BD, Lovern MR, Stockis A, Sargentini-Maier ML, Karlsson MO, Friberg LE (2009) A pharmacodynamic Markov mixed-effects model for determining the effect of exposure to certolizumab pegol on the ACR20 score in patients with rheumatoid arthritis. *Clin Pharmacol Ther* 86:387–395
5. Keystone EC, Genovese MC, Klareskog L, Hsia EC, Hall ST, Miranda PC, Pazdur J, Bae SC, Palmer W, Zrubek J, Wiekowski M, Visvanathan S, Wu Z, Rahman MU (2009) Golimumab, a human antibody to tumour necrosis factor α given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. *Ann Rheum Dis* 68:789–796
6. Simponi: Package insert (2009) Centocor Ortho Biotech, Inc., Malvern, PA
7. Xu Z, Vu T, Lee H, Hu C, Ling J, Yan H, Baker D, Beutler A, Pendley C, Wagner C, Davis HM, Zhou H (2009) Population pharmacokinetics of golimumab, an anti-tumor necrosis factor- α human monoclonal antibody, in patients with psoriatic arthritis. *J Clin Pharmacol* 49:1056–1070
8. Beal SL (2005) Conditioning on certain random events associated with statistical variability in PK/PD. *J Pharmacokinet Pharmacodyn* 32:213–243
9. Hu C, Xu Z, Zhang Y, Rahman MU, Davis HM, Zhou H (2010) A population approach for exposure-response modeling of golimumab in patients with rheumatoid arthritis. *J Clin Pharmacol* (in press)
10. Hu C, Zhou H (2008) An improved approach for confirmatory phase III population pharmacokinetic analysis. *J Clin Pharmacol* 48:812–822
11. Zhang L, Beal SL, Sheiner LB (2003) Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *J Pharmacokinet Pharmacodyn* 30:387–404
12. Freedman D, Pisani R, Purves R (1998) *Statistics*, 3rd edn. Norton, New York