

Bounded outcome score modeling: application to treating psoriasis with ustekinumab

Chuanpu Hu · Newman Yeilding · Hugh M. Davis ·
Honghui Zhou

Received: 29 March 2011 / Accepted: 4 June 2011 / Published online: 18 June 2011
© Springer Science+Business Media, LLC 2011

Abstract Disease status is often measured with bounded outcome scores (BOS) which report a discrete set of values on a finite range. The distribution of BOS data is often non-standard, e.g., J- or U-shaped, thus making standard analysis methods that assume normality inappropriate. Data transformations aiming to achieve normality with BOS can be much more difficult than with many other types of skewed distributions, and application of methodologies explicitly dealing with this problem has not been previously published in pharmacokinetic/pharmacodynamic modeling literature. In this analysis, a coarsened latent variable (CO) approach is augmented with flexible transformations and applied for the purpose of demonstrating ustekinumab effects on four clinical components (involved body surface area, induration, erythema, and scaling) in patients with moderate to severe psoriasis from two Phase 3 studies. Patients were randomized to receive placebo or ustekinumab 45 or 90 mg, followed by randomized withdrawal and long-term extension periods. The approach was used together with a previously established novel semi-mechanistic, mixed-effect exposure-response model integrated with placebo effect and disease progression, and with potential influence of dropout investigated. An additional transformation further modifying both tails of the standard logit transformation in the original CO approach was shown to be necessary in this application.

Keywords Monoclonal antibody · Box–Cox transformation · Probit regression · Informative censoring · Nonlinear mixed effect modeling · NONMEM

C. Hu (✉) · H. M. Davis · H. Zhou
Pharmacokinetics and Pharmacometrics, Biologics Clinical Pharmacology, Centocor R&D,
A Division of Johnson & Johnson Pharmaceutical Research & Development, LLC.,
200 Great Valley Parkway, Malvern, PA 19355, USA
e-mail: chu25@its.jnj.com

N. Yeilding
Clinical Immunology, Centocor R&D, A Division of Johnson & Johnson Pharmaceutical Research
& Development, LLC., Malvern, PA, USA

Introduction

Psoriasis is a chronic immune-mediated skin disorder, and interleukin (IL)-12 and -23 have been implicated in the pathogenesis of this disease [1–3]. Ustekinumab is a human IgG₁ monoclonal antibody that binds with high specificity and affinity to the shared p40 subunit of IL-12 and IL-23 and blocks interaction with the IL-12Rβ1 cell surface receptor. Data were obtained from two large Phase 3 trials of ustekinumab in patients with moderate to severe psoriasis (PHOENIX 1 and PHOENIX 2) [4, 5]. Previously reported analyses of these data include population pharmacokinetics (PK) of ustekinumab in patients with psoriasis [6, 7], population pharmacokinetic/pharmacodynamic (PK/PD) modeling of improvement in disease conditions using a mechanistic approach using Psoriasis Area and Severity Index (PASI) scores as a continuous measure [8], and a novel semi-mechanistic approach for ordered categorical data in conjunction with disease progression and informative dropout using Physician's Global Assessment (PGA) scores [9]. As noted previously [9], it is of clinical interest to determine whether ustekinumab affects any or all of the four clinical components (involved BSA, erythema, induration and scaling) comprised in the PASI score, and that is the focus of this manuscript.

The PASI score is a composite endpoint that reflects the clinical features of psoriasis, ranges from 0 to 72, and is calculated as follows. First, the body is divided into four regions: the head (h), trunk (t), upper extremities (u), and lower extremities (l), which account for 10, 30, 20, and 40%, respectively, of the total body surface area (BSA). Each of these areas is assessed separately for erythema, induration and scaling, which are each rated on a scale of 0–4 (0 = none, 1 = slight, 2 = moderate, 3 = severe, and 4 = very severe). Furthermore, a 7-point scale is used to assess the percentages of involvement in each of the four areas (0 = no involvement, 1 = 1–9%, 2 = 10–29%, 3 = 30–49%, 4 = 50–69%, 5 = 70–89%, and 6 = 90–100%). The PASI formula is given by

$$\text{PASI} = 0.1(E_h + I_h + S_h)A_h + 0.3(E_t + I_t + S_t)A_t + 0.2(E_u + I_u + S_u)A_u + 0.4(E_l + I_l + S_l)A_l,$$

where E, I, and S correspond to erythema, induration, and scaling, respectively, and A corresponds to percentages of involved BSA [10].

As noted previously [9], it is of clinical interest to determine whether ustekinumab affects any or all of the four clinical components (involved BSA [iBSA], erythema, induration and scaling) that compose the PASI score using the same weights. That is, for the overall weighted measure for involved body surface area,

$$i\text{BSA} = 0.1A_h + 0.3A_t + 0.2A_u + 0.4A_l,$$

and similarly, Erythema = $0.1E_h + 0.3E_t + 0.2E_u + 0.4E_l$, Induration = $0.1I_h + 0.3I_t + 0.2I_u + 0.4I_l$, and Scaling = $0.1S_h + 0.3S_t + 0.2S_u + 0.4S_l$. These disease component measures take the form of bounded outcome scores (BOS) with a discrete set of values from a restricted interval with natural boundaries. Specifically, possible iBSA values range from 0 to 6, and erythema, induration, and scaling values range from 0 to 4, all with increments of 0.1. Note that values outside these boundaries are neither possible nor meaningful.

BOS outcomes are also ordered categorical variables. When the number of possible response categories is small (e.g., <5), the standard analysis approach is logistic regression. When the number of categories is large (e.g., >10), the BOS outcomes are most often analyzed as usual continuous outcomes. This requires the assumption that the scores are normally distributed, which, if violated, can lead to biased analysis results [11, 12]. In principle, the distribution of BOS depends on the granularity of its construction. A coarse scale on one or both ends may result in (1) skewed distribution; and/or (2) substantial amount of data at the boundaries making the distribution J- or U-shaped. Both cases lead to violation of normality. Note that J-shaped data imply skewness, but the reverse is not true as data could be distributed more on one side and still attenuate at the boundary.

Ordinary skewness of condition (1) can typically be handled effectively by data transformation. For example, log-transformation is often used for positive variables. Another standard choice, in the case of bounded outcomes, is to first transform the data onto the interval (0, 1) and then use the logit transformation $\text{logit}(x) = \log[x/(1 - x)]$. However, the boundary-attaining aspect of condition (2) creates difficulty for the usual data transformation approach. For example, at the boundary points 0 and 1, the logit function approaches infinity, thus preventing its use. One might think that this problem could be avoided by first transforming the BOS values to a slightly smaller interval of $[\delta, 1 - \delta]$ with $\delta > 0$ and then applying the logit transformation. However, the choice of δ will have to be arbitrary and the results, especially with J- or U-shaped distributions, are likely to be sensitive to the choice, thus making the results difficult to interpret. This has been pointed out in statistical literature, most recently by Hutmacher et al. [13]. Based on an idea similar to a censored-data approach to handling PK data below the quantification limit [14], they proposed a transformation approach using the Aranda-Ordaz asymmetric family [15] for the analysis of continuous BOS data.

Recently, a conceptually different and appealing general approach has been proposed in the statistical literature [16, 17]. It presumes a latent (unobserved) underlying continuous variable of which the observed BOS is a “coarsened” version. This coarsened latent variable (CO) approach works as follows. The original BOS value can be easily standardized onto the closed interval [0, 1] by a linear transformation. Let m be the number of possible BOS values taking the form of k/m , where $k = 0, 1, \dots, m$. A latent variable U , taking the value on the open interval (0, 1), is related to the BOS outcome through:

$$\text{BOS} = \frac{k}{m} \text{ if and only if } a_k \leq U < a_{k+1}, \quad \text{for } k = 0, \dots, m$$

where $a_k = (k - 0.5)/m$ and $a_{k+1} = (k + 0.5)/m$, with $a_0 = 0$ and $a_{m+1} = 1$. It is also assumed that:

$$\text{logit}(U) = p + \varepsilon \tag{1}$$

where p is the model prediction on the transformed scale, and ε is normally distributed with mean 0 and variance σ^2 . Then, the likelihood of observing $\text{BOS} = k/m$, conditional on its individual prediction p , is calculated as:

$$\Phi\left(\frac{z_k^{(u)} - p}{\sigma}\right) - \Phi\left(\frac{z_k^{(l)} - p}{\sigma}\right) \quad (2)$$

where $z_k^{(l)} = \text{logit}(a_k)$, $z_k^{(u)} = \text{logit}(a_{k+1})$, and Φ is the cumulative distribution function of the standard normal distribution. To formulate the mixed-effect model likelihood, let $p = p(\theta, \eta)$, where θ represents model fixed-effect parameters, $\eta \sim N(0, \Omega)$ represents between-subject variability distributed with mean 0 and variance–covariance matrix Ω , and $\mathbf{y}_i = \{y_{i,1}, \dots, y_{i,n_i}\}$ is the observed data vector of subject i where $y_{i,j}$ is the j th observation having the form of k/m . Then, the marginal likelihood of subject i is given by

$$L(\theta, \Omega, \sigma^2 | y_i) = \int \prod_j^{n_i} l(y_{i,j} | \eta) f(\eta) d\eta \quad (3)$$

where $l(y_{i,j} | \eta)$ is given by Eq. 2, and $f(\eta)$ is the multivariate normal density function of $N(0, \Omega)$. The overall likelihood is the product of the marginal likelihood over all subjects.

The CO approach may be viewed as a standard framework to set the appropriate residual error model for BOS outcomes when the granularity is coarse enough to prevent the outcome from being treated effectively as continuous. The use of normal distribution requires that the logit transformation of the latent variable U yields a symmetric residual distribution. As this is not guaranteed, additional transformation in addition to the logit transformation may be necessary. In principle, additional transformations under the logit transformation may be viewed as providing flexibility in the residual error model. Such transformations have been used in the statistical literature on logistic regression under the term of link functions. This motivated the use of link functions with the BOS analysis, as the CO approach closely resembles logistic regression.

In addition to the methodological objective of investigating the CO approach in PK/PD modeling, the clinical objective of this analysis was to determine whether treatment with ustekinumab improved iBSA, erythema, induration, and scaling, the need of which was noted earlier [9]. The underlying PK/PD model was based on a previously established model based on PGA scores, which were summary measures for three of the four clinical components considered here. As this approach is relatively new, an external validation was used to evaluate the model built for iBSA.

Methods

Clinical studies

The patient populations and study designs for the randomized, double-blind, placebo-controlled PHOENIX 1 and PHOENIX 2 trials were complex and details have been previously described [9]. Briefly, in PHOENIX 1, a total of 766 patients were assigned to receive subcutaneous injections of ustekinumab 45 mg,

ustekinumab 90 mg, or placebo at weeks 0 and 4, followed by active treatment every 12 weeks or placebo crossover to ustekinumab 45 or 90 mg starting at week 12 (weeks 12–40), followed by randomized withdrawal and long-term extension periods [4]. In PHOENIX 2, a total of 1,230 patients were assigned to the same treatment groups and design until week 28, followed by dose schedule optimization and long-term extension periods [5]. All data collected up to week 100 in both trials were included in the current analyses. In PHOENIX 1, most patients had data up to week 136, and 157 patients had data at or beyond week 152.

Serum ustekinumab measurement

Blood samples for the measurement of serum ustekinumab concentrations were collected at weeks 0, 4, 12, 16, 24, 28, 40, 44, 48, 52, 56, 60, 64, 68, 72, 76, and 88 (and every 24 weeks thereafter, if available) in PHOENIX 1 and at weeks 0, 4, 12, 16, 20, 24, 28, 40, 52, and 88 (and every 24 weeks thereafter, if available) in PHOENIX 2.

Disease score measurement

PASI scores were collected at weeks 0, 2, and 4, and every 4 weeks thereafter up to week 88 or study unblinding, then at week 100 and every 12 weeks thereafter in PHOENIX 1 and at weeks 0, 2, and 4, and every 4 weeks thereafter up to week 52 or study unblinding, and every 12 weeks after week 52 in PHOENIX 2.

Dataset for PK/PD modeling and validation

Only patients with available PK data were included in the dataset. For iBSA modeling, data from PHOENIX 2 were used for model building, and data from PHOENIX 1 were reserved for model validation. All PK and disease score measurement data through week 100 were included in the current analysis. For PHOENIX 2, there were 9,723 PK observations and 21,711 disease score measurements from a total of 1,230 patients. For PHOENIX 1, there were 9,617 PK observations and disease score measurements from a total of 765 patients. The modeling of erythema, induration, and scaling used the combined PHOENIX 1 and PHOENIX 2 dataset. In the PHOENIX 2 study, 211 (17.2%) patients discontinued the study at various times before week 100; the exact discontinuation times were available for 68 (5.5%) patients. In PHOENIX 1, 162 (21.2%) patients discontinued the study before week 152, and the discontinuation date was known for 48 (6.3%) patients.

Alternative motivation of coarsened latent variable approach and link function

As mentioned earlier, BOS outcomes are ordered categorical variables, for which logistic regression is applicable, though the number of intercept terms may be large. A similar and equally viable approach is probit regression [18], which uses the normal cumulative distribution function, Φ , in place of the logit function as the link

between the predictor and the cumulative probabilities of achieving the response values. The use of the CO approach can be motivated by using logit-transformed, fixed intercepts in probit regression [17]. Specifically, BOS outcomes may be viewed as interval censored observations, where the intervals are given by the cut points, a_k , preceding Eq. 1 as on the original scale. The a_k are then logit-transformed to z_k in Eq. 2 as values on the real line, which are then used for probit modeling. Because a_k are fixed, an additional scaling-like term, σ , can be estimated. It would not be identifiable in probit regression due to confounding with the estimation of intercepts. This allows the CO approach to be viewed as using σ as a parsimonious alternative to estimating all intercept parameters.

On the other hand, with all cut points fixed in the CO approach, the scaling term σ alone may not provide appropriate values for the assumption of normality. In addition, it may also lack the flexibility of the ordinal (ordered categorical, probit/logistic) regression approaches. This suggests the need of flexible transformation families, in a sense to further map the cut points into more appropriate values, and also making the normality assumption tenable.

Using link functions with BOS analysis

The link function approach as used in logistic regression [19] works as follows. Let U be the PK/PD model predicted mean outcome on the interval $(0, 1)$, and $p(\theta, \eta)$ be the model predictor for the outcome where (θ, η) represent the set of model fixed-effect parameters and between-subject random effects. The link function g is defined through

$$g(U) = p(\theta, \eta)$$

In particular, choosing the function $g(x) = \log[x/(1-x)]$ gives the standard logistic regression.

More general link functions are defined through families of functions with extra parameters allowing more flexibility. This is related to and may be motivated by transformations [18]. The often-used Box–Cox transformation $(y^\lambda - 1)/\lambda$, defined for $y \geq 0$ [20], is essentially a power transformation with parameter λ determining the skewness of the transformation near 0. This serves to motivate the Aranda-Ordaz asymmetric family, a commonly used link function family in logistic regression, defined by

$$g(x, \lambda) = \log \left(\frac{(1-x)^{-\lambda} - 1}{\lambda} \right)$$

where λ is a parameter to be estimated. It reduces to the logit link when $\lambda = 1$. A more complex link function family also used in logistic regression is the Czado family [21], with two parameters (λ_1, λ_2) , defined by

$$h(x, \lambda_1, \lambda_2) = \begin{cases} \frac{(x+1)^{\lambda_1} - 1}{\lambda_1} & \text{if } x \geq 0 \\ -\frac{(-x+1)^{\lambda_2} - 1}{\lambda_2} & \text{if } x < 0 \end{cases}$$

and $g(x, \lambda_1, \lambda_2) = h^{-1} [\text{logit}(x), \lambda_1, \lambda_2]$. It reduces to the logit link when $\lambda_1 = \lambda_2 = 1$. It can be seen that the Aranda-Ordaz asymmetric family focuses on the left tail of the distribution, but can also be easily modified, through the transformation $x \rightarrow (1 - x)$, to focus on the right tail instead. The Czado family modifies both tails.

The link function approach can be adapted to the CO approach for BOS analysis by simply taking the transformation of U as the latent variable in Eq. 1, instead of the probability of the binary outcome variable in logistic regression. This leads to replacing the logit transformation in Eqs. 1–3 with one of the transformation families above. More specifically, in Eq. 2, $z_k^{(l)} = g(a_k)$ and $z_k^{(u)} = g(a_{k+1})$ are used with parameter λ or (λ_1, λ_2) inside the link function g to be estimated, where the choice of parameter depends on whether the Aranda-Ordaz or Czado link function is used. In this setting, the parameters to be estimated $(\theta, \Omega, \sigma^2)$, are augmented with λ or (λ_1, λ_2) to the left side of Eq. 3. Because the transformation parameters λ or (λ_1, λ_2) modify the scale of the predictor p on which σ operates, estimates of λ or (λ_1, λ_2) and σ are expected to be highly correlated. This aspect is similar to analyzing continuous data with Box–Cox transformation.

As mentioned above, the cut points in the CO approach are related to intercept parameters in ordinal regression modeling. The difference is that the number of estimated parameters does not increase with the number of BOS levels with the CO approach, in which the use of the Aranda-Ordaz and Czado transformations may be viewed as a parsimonious intermediary. It can thus be seen that ordinal regression would be appropriate when the number of BOS levels is small (e.g., <5), and the benefit of parsimony can be achieved by using the CO-transformation approach when the number of BOS levels is high (e.g., >10), allowing sufficient information to estimate the transformation parameters. General benefits from parsimony include estimation efficiency, stability, and run time savings. For example, there may be few observations at certain BOS levels, leading to difficulty in estimating the corresponding intercepts with ordinal regressions. The standard way to handle this in ordinal regressions is to merge those levels. However, as the number of BOS levels increases, the potential levels that need to be merged tend to become more difficult to determine, making the analysis less stable and more time consuming.

At a conceptual level, the analysis framework should appropriately reflect the nature of the data. That is, the use of the cut-value of 0.5 on the original data scale (in $a_k = (k - 0.5)/m$) should indeed reflect how BOS scores are obtained. In many cases, this is a result of rounding [16]. However, in this application, instead of arising from a single assessment, the data were constructed as weighted averages of multiple scales, thus one might question whether 0.5 is still the appropriate cut-value. This can be seen from the fact that, as each individual scale is an integer, it is reasonable to assume that the assessment incurs a rounding error of 0.5. As each scale incurs this rounding error, this is also the rounding error for the weighted averages, i.e., the BOS values.

The CO framework, with or without transformations, essentially considers BOS outcomes as interval-censored data, where the intervals are determined by the cut points, a_k . Therefore, diagnostics are more difficult than continuous data. Formally, the link functions transform the predictor, not the data; therefore, determining the

appropriate transformation is complicated. The departure of raw data distribution from normality may suggest the need for transformations, although it is difficult to know which one is the best. In principle, examining transformations of the raw data distribution may be most useful when the number of BOS levels is large and the scores are not close to the boundaries, which make BOS outcomes more like continuous data, but then the CO approach would not be necessary. The presence of mixed-effects adds further complexities to this issue.

Model evaluation

Model diagnosis under the CO approach can be a challenge due to the nature of BOS outcomes as interval censored data. Modified Cox-Snell residuals were used with dropout modeling in previous applications [9, 22], however they do not account for within-subject correlation and, thus, would not be appropriate in this setting. Because the interest in modeling typically lies in appropriately predicting the distribution of data, a visual predictive check (VPC), which compares the observed data distribution with the predicted, was used.

PK/PD model

The semi-mechanistic model described by Hu et al. [9] was adopted as follows. The population PK model and the empirical Bayesian estimates of the individual PK parameters, based on a confirmatory analysis using an established one-compartment model with first-order absorption [6], were retained. For the exposure–response portion of the model, the intercept term in the logistic regression was simply replaced with a likewise baseline term. Then, the model prediction in Eq. 1 is given by:

$$p = b + f_z(t) + f_p(t) + f_d(t) + \eta \quad (4)$$

where b is baseline, $f_z(t)$ is disease progression, $f_p(t)$ is placebo effect, $f_d(t)$ is drug effect, and η is a normally distributed random variable with a mean of 0 and represents the between-subject variability. Specifically, the disease progression was modeled as:

$$f_z(t) = \beta t \quad (5)$$

where β represents the rate of decline. The placebo effect was modeled empirically as:

$$f_p(t) = \text{Plb}_{\max} [1 - \exp(-R_p t)] \quad (6)$$

where Plb_{\max} is the maximum placebo effect and R_p is the rate of onset. The drug effect was modeled using a latent variable $R(t)$, governed by:

$$\frac{dR(t)}{dt} = k_{in} \left(1 - \frac{C_p}{IC_{50} + C_p} \right) - k_{out} R(t) \quad (7)$$

where C_p is the ustekinumab concentration, and k_{in} , IC_{50} , and k_{out} are parameters in a Type I indirect response model. It was further assumed that $R = 1$ at baseline,

i.e., $R(0) = 1$, leading to $k_{in} = k_{out}$. The reduction of $R(t)$ was assumed to drive the drug effect through:

$$f_d(t) = DE(1 - R(t)) \quad (8)$$

where DE is the drug effect. Eqs. 4–8 along with Eq. 1 (or equivalently, Eq. 2) establish the link between ustekinumab concentrations and the BOS scores in patients with psoriasis.

The link between BOS and ordered categorical data allows informative dropout modeling as implemented in Hu et al. [9] to be adopted straightforwardly. The details of the methodology are out of the scope of this manuscript.

Implementation

NONMEM Version 7 [23] was used for model development. The Laplacian option was used with a sequential estimation approach, and simultaneous fitting [24] was not performed due to the long run time, with a single model run taking approximately 15 h on an IBM IntelliStation Z Pro Workstation. Because nearly ten concentration measurements were available per patient, the sequential PK/PD method was expected to perform satisfactorily. This assumption was supported by previous applications using different endpoints collected from the same study [8, 9]. Hypothesis tests for drug effect were conducted with the likelihood ratio test, for which a decrease in the NONMEM minimum objective function of 10.83, corresponding to a nominal p -value of 0.001, was considered the threshold criterion for statistical significance.

Results

As shown by Zhou et al. [8], the baseline demographic and disease characteristics for patients included in the PK analysis were similar in PHOENIX 1 and PHOENIX 2. Disease score distributions at baseline and weeks 20, 40, and 80 are shown in Fig. 1. The right boundaries (6 for iBSA and 4 for all others) were attained at baseline, and the left boundary 0 was attained at later weeks. Distributions of scores at baseline were close to symmetric, which in part reflect trial entry criteria. Distributions at later weeks were reversely J-shaped. The distributions appeared to shift marginally further to the left from week 20 to week 80, suggesting treatment benefit over those times. As noted earlier, it is difficult to determine the best transformations solely based on these visual examinations; however, they do suggest that transformations may be needed in analyzing these data.

iBSA analysis

Initial model with PHOENIX 2 data

Applying the CO approach with logit transformation to the data from PHOENIX 2 had a likelihood ratio test for the drug effect that was strongly significant, with a

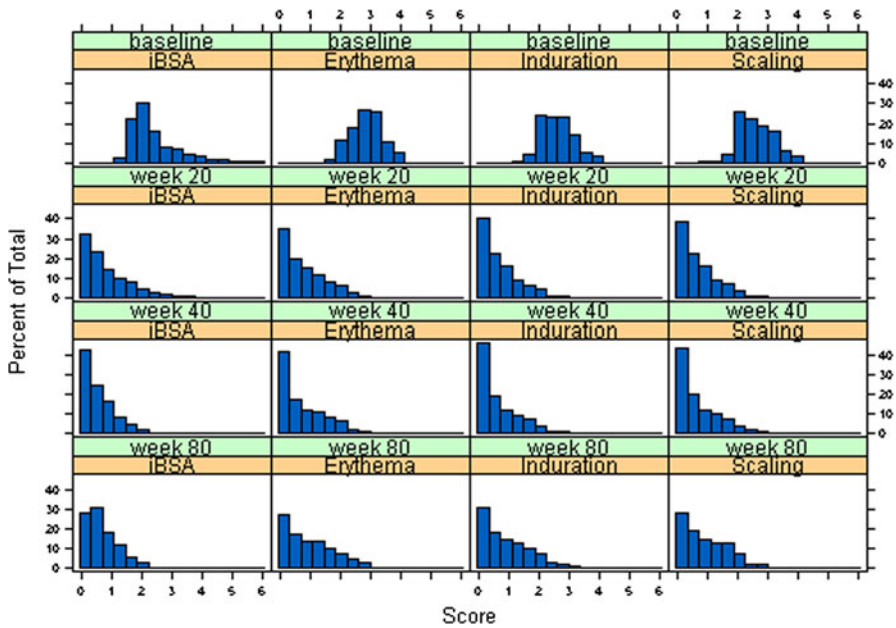


Fig. 1 Disease score distributions at baseline and week 40 for iBSA, erythema, induration and scaling

NONMEM objective function change $>5,000$. Figure 2 shows the VPC result, which clearly demonstrated a therapeutic effect of ustekinumab. Despite the notable consistent downward bias, the model described the temporal changes well. This demonstrated that logit transformation led to a distribution skewed to the left, consistent with the results shown in Fig. 1. Consequently, additional transformations were investigated.

The Aranda-Ordaz transformation was considered next with the purpose of correcting the left skewness. Estimation was difficult, with NONMEM consistently terminating with the message “PROXIMITY OF LAST ITERATION EST. TO A VALUE AT WHICH THE OBJ. FUNC. IS INFINITE.” However, VPC results of some initial analyses (not shown) suggested that the transformation corrected the bias, which supported the need for additional transformations. Further investigation of selected individual likelihood profiles confirmed the high nonlinearity of the likelihood surface, and showed that termination was caused by certain large observations. As the Aranda-Ordaz transformation emphasizes the left tail, this suggested the need for transformations modifying both the left and right tails.

The Czado transformation was then considered. Estimation was stable, and the results are given in Table 1. Standard errors of parameter estimates were generally an order of magnitude smaller than the parameter estimates. The NONMEM objective function value decreased from the original logit transformation model by more than 9,000. Furthermore, the Czado transformation parameters (λ_1 , λ_2) had estimates >3 and well-bounded away from 1 in consideration of their standard errors, confirming significant departure from the logit transformation. Estimation

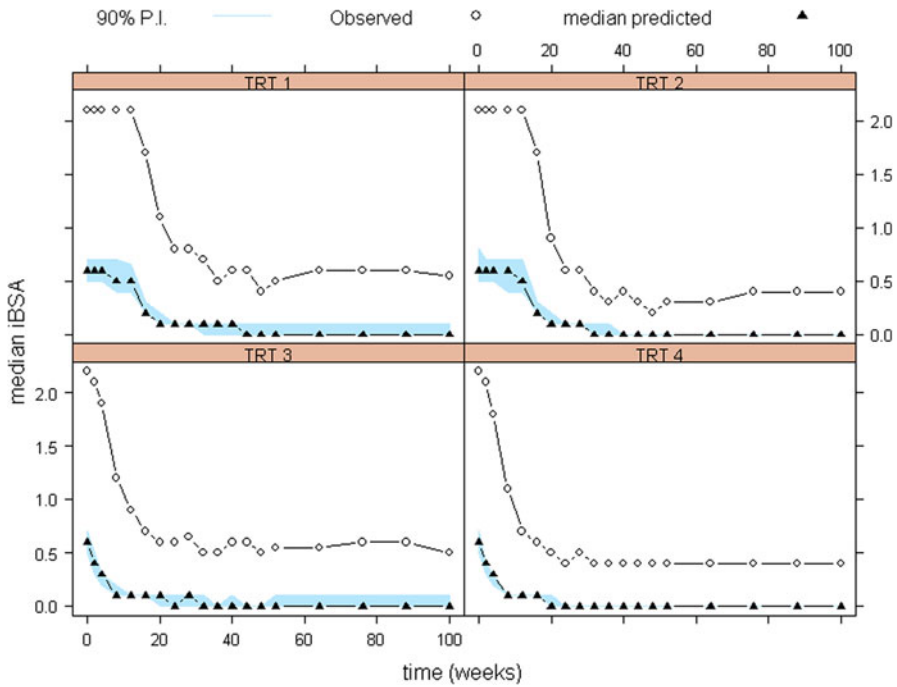


Fig. 2 VPC of the initial iBSA PK/PD model using logit transformation built with PHOENIX 2 data. Observed data medians and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

was stable, with standard errors generally a magnitude smaller than parameter estimates, which is to be expected in this data-rich situation with a large number (61) of BOS levels. As expected, estimates of σ , λ_1 , and λ_2 were highly correlated, with absolute values ranging from 0.71 to 0.94. The condition number (306), calculated as the ratio between the largest and smallest eigenvalues of the correlation matrix, was also reasonable. VPC results in Fig. 3 show that the overall bias was corrected. Discrepancies between the observed and model predictions of reduced magnitude still remained, with an over-prediction at the placebo periods and an under-prediction at later periods. A detailed investigation suggested that these discrepancies were likely caused by structural and/or between-subject variability model misspecification and were not due to the residual error model or link function. Additional between-subject random effects on parameters other than baseline were attempted but consistently resulted in abnormal termination in NONMEM. This difficulty apparently resulted from the LAPLACIAN approximation used in this circumstance; limited attempts with the Importance Sampling (METHOD = IMP) option seemed to resolve the problem. However the long computation time (30 h for two between-subject random effects on a workstation with 3.47 GHz CPU) prevented the option from being used for extensive exploration. This did not appear to affect the mean predictions and hence the

Table 1 Model parameter estimates

Parameter	Estimate (% RSE)				
	iBSA-initial	iBSA-final	Erythema	Induration	Scaling
σ	0.200 (4.5)	0.204 (3.5)	0.373 (1.9)	0.345 (2.0)	0.347 (2.0)
b	-0.222 (3.5)	-0.212 (3.0)	0.585 (2.5)	0.380 (3.4)	0.359 (3.5)
plb_{max}	0.109 (27.2)	0.221 (8.6)	0.408 (7.1)	0.387 (7.6)	0.354 (6.9)
R_p (day ⁻¹)	0.0118 (14.7)	0.0111 (6.8)	0.0177 (7.1)	0.0185 (7.2)	0.0215 (8.6)
β (day ⁻¹)	0.000182 (13.4)	0.000114 (11.3)	0.000217 (10.1)	0.000214 (9.8)	0.000192 (9.6)
k_{out} (day ⁻¹)	0.0118 (4.0)	0.0123 (2.9)	0.0203 (3.5)	0.0220 (6.1)	0.0222 (3.6)
DE	0.813 (5.5)	0.705 (4.5)	1.33 (3.1)	1.15 (3.9)	1.13 (3.0)
λ_1	9.71 (11.1)	9.2 (9.2)	2.65 (3.3)	3.16 (3.8)	3.32 (3.7)
λ_2	3.83 (3.5)	3.75 (2.8)	3.55 (1.9)	3.68 (1.9)	3.57 (1.9)
$Var(\eta)$	0.0587 (9.9)	0.0627 (8.1)	0.159 (4.5)	0.123 (4.9)	0.133 (4.8)

σ standard deviation of latent variable variability, b intercept, plb_{max} maximum placebo effect, R_p rate of placebo effect onset, β rate of disease progression, k_{out} disease amelioration rate, DE drug effect, λ_1 Czado transformation parameter modifying right tail, λ_2 Czado transformation parameter modifying left tail, $Var(\eta)$ variance of inter-subject variability η , RSE relative standard error

clinical objective of this analysis, which is to demonstrate the ustekinumab effect along with the CO approach.

Modeling of dropout showed that it was neither completely random nor informative, but random in the sense of Hu et al [9]. This difference from previous results [9] suggests that iBSA may be a better predictor of dropout than PGA.

External validation of initial model with PHOENIX 1 data

Figure 4 shows the VPC results of applying the initial model with Czado transformation to the PHOENIX I data. In this analysis, the VPC was used to provide a qualitative overview of the results. The overall results appeared to be consistent with the model building results shown in Fig. 3. The predictive ability of the model was sufficient to address the question of clinical interest on whether ustekinumab could improve iBSA.

Final model with combined PHOENIX 1 and 2 data

The initial model with Czado transformation was fitted to the combined data, and the results are given in Table 1. Parameter estimates were generally similar to the initial ones, and the standard errors decreased as the result of incorporating more data. Correlation of estimates of σ , λ_1 , and λ_2 remained similar, with absolute values ranging from 0.72 to 0.94. The condition number was 211, which was reasonable. Figure 5 shows the VPC results of the final model, with observed and predicted distributions of the 5th and 95th percentiles included in order to examine how well the variability was predicted. Results of observed scores at weeks 0, 20, 40, and 80

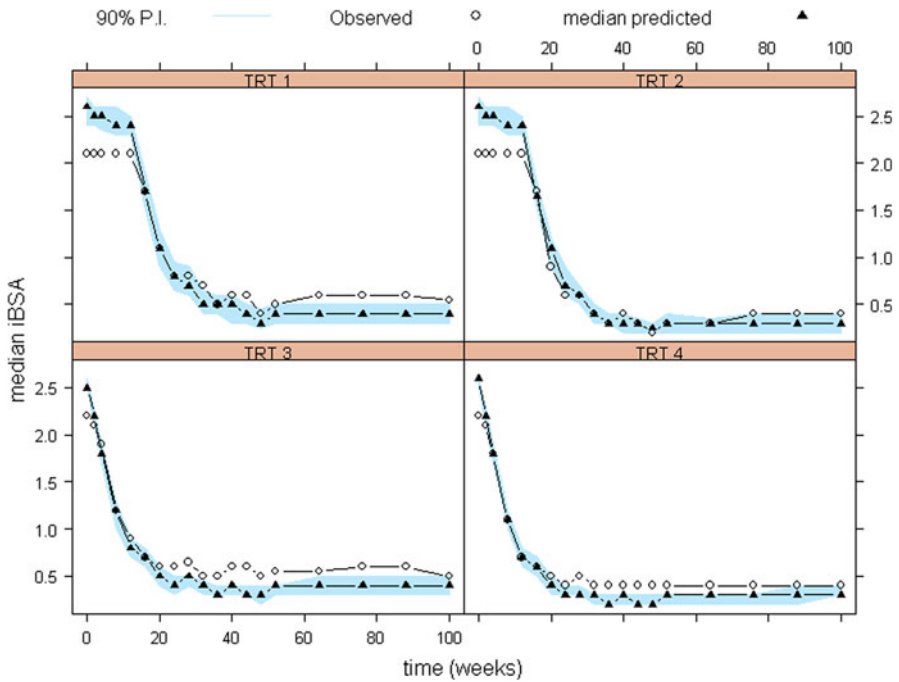


Fig. 3 VPC of the initial iBSA PK/PD model using Czado transformation built with PHOENIX 2 data. Observed data medians and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

may be viewed as alternative looks of distributions shown in Fig. 1. VPC of the median appeared to be consistent with those shown in Figs. 3 and 4. VPC of the 95th percentile appeared to over-predict the observed data from week 40 to 80 onwards, suggesting possible improvement in between-subject variability modeling.

As a comparison, the full logistic regression model with 60 intercept terms was also attempted. The random effect term was fixed to zero, in part due to the long array lengths that resulted from NONMEM dictating the use of the “\$ABBREVIATED DERIV2 = NO” option, preventing the use of the LAPLACIAN estimation option. The computation time was over 30 times longer than the CO transformation approach also with the random effect term fixed to zero. Initial results appeared to show consistency with the CO approach; however, they were highly sensitive to initial parameter estimates, and efforts to assure convergence could not be completed due to the long computation time.

Erythema, induration, and scaling analyses

With an effort to maintain the confirmative nature of the analysis to the extent possible, the final iBSA model with Czado transformation was retained for the remaining analyses. Likelihood ratio tests for the drug effect were significant with

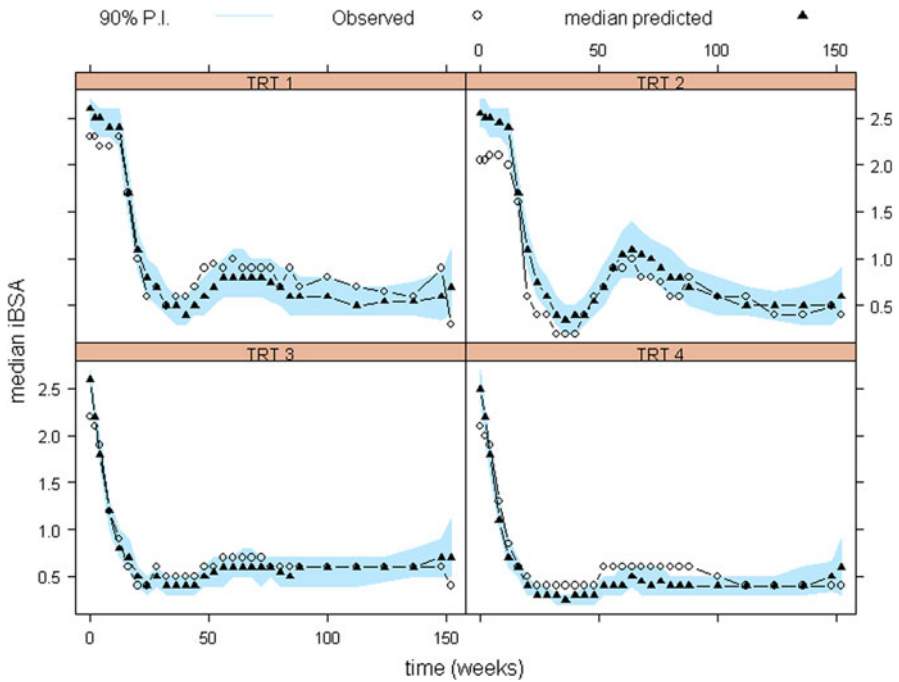


Fig. 4 External-validation VPC using PHOENIX 1 data for the initial iBSA PK/PD model using Czado transformation. *TRT*, treatment groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 mg or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

NONMEM objective function changes $>10,000$. All estimations were stable, and the results are given in Table 1. Correlation of estimates of σ , λ_1 , and λ_2 were slightly reduced compared with iBSA, with absolute values ranging from 0.67 to 0.85. The condition numbers were approximately 100–200, which were reasonable. The Czado transformation parameters (λ_1 , λ_2) were well-bounded away from 1 in consideration of their standard errors, indicating departures from the logit transformation. Parameter estimates were largely similar, suggesting that ustekinumab improved these disease measures in similar ways. These estimates were difficult to compare with the iBSA estimates, because iBSA had a different range (0–6) than erythema, induration, and scaling (0–4). VPC results are shown in Figs. 6, 7, and 8. The results appeared relatively similar and indicated the presence of a therapeutic effect from ustekinumab therapy. In contrast to iBSA results in Fig. 5, VPC of the 95th percentile now appeared to be reasonable.

Discussion

Disease score measurements in the form of BOS take a discrete set of values, with a substantial amount on one or both boundaries of closed intervals, leading to violation of the normality assumption as shown in Fig. 1. This provides a challenge to

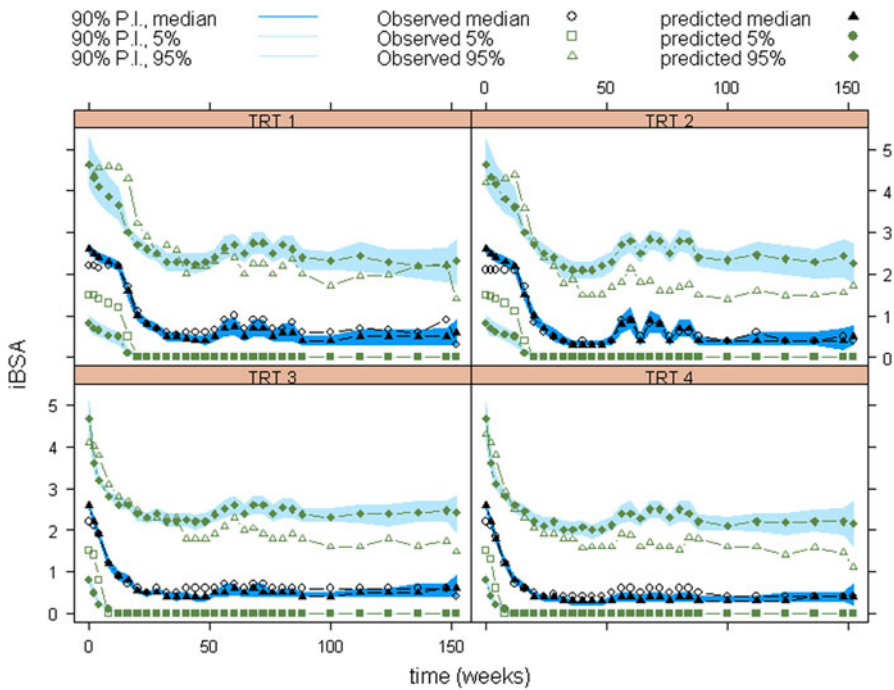


Fig. 5 VPC of the final iBSA PK/PD model using Czado transformation, built with combined data from PHOENIX 1 and 2. Observed data medians and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

analyzing BOS appropriately, and the statistical methodology is still evolving. Hutmacher et al. [13] noted that predicting only feasible outcome values has the advantages of providing interpretability and assuring valid predictions and predictive checks for various quantities of practical interest, e.g., quantiles or prediction intervals (PI) of the outcomes, compared with the often used approach of simply treating BOS as ordinary continuous outcomes. Their approach is suitable for BOS having many possible outcome categories and thus treated as continuous variables attaining the boundary. Despite their statement that the method is related to those developed for data with only a lower bound, the approach could be extended to data with values of either boundary simultaneously. Another approach that has been used for general BOS outcomes is Beta-regression [25], however an artificial correction factor has to be used for observations at the boundary, which may be viewed as a shortcoming. This effect of the correction factor is more notable especially when a substantial amount of observations are at the boundary. The CO approach provides an appealing statistical framework that first postulates that BOS outcomes are caused by an underlying latent variable, and then translates the latent variable onto a scale where normality is suitable. It is applicable for general BOS outcomes, even if the number of categories of the outcome (e.g., say 5–10) is too small to be treated as continuous but too large to be analyzed as ordered categorical data.

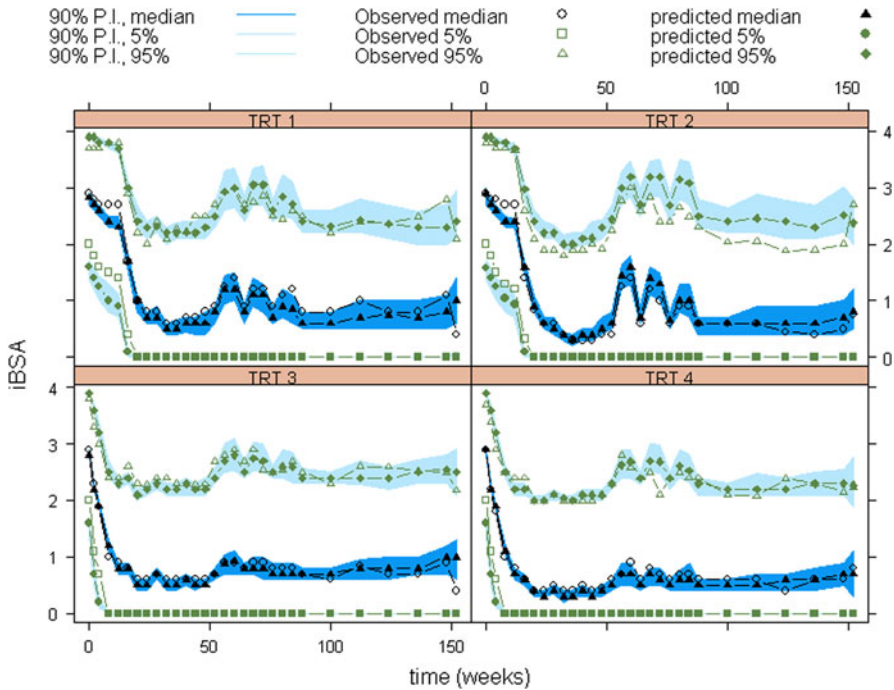


Fig. 6 VPC of the erythema pharmacokinetic/pharmacodynamic model. Observed data medians and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 mg or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

The current analysis presents the first application of the CO approach in PK/PD modeling and with augmentation of some general link functions used in logistic regression. As illustrated in Fig. 2, the standard logit transformation alone in the original CO approach may not be appropriate. In addition, the Aranda-Ordaz transformation focusing on only the left tail was not sufficient. The more general Czado transformation with two parameters modifying both tails was shown to be necessary in this application. As the Czado transformation closely relates to the Box–Cox transformation, this finding appears to provide evidence contrary to the intuitive view presented by Lesaffre et al. [16] that the Box–Cox transformations were unlikely to be useful in clinical trial settings. It is noted however that this application involved a large dataset consisting of two large phase III clinical trials; and such transformations may not be supportable with small datasets having either few categories or observed values near the boundaries.

BOS are also ordered categorical data, and the CO approach with transformation may be viewed as a parsimonious alternative to ordinal regression. The efficiency gains and computational savings can be significant when the number of possible response levels is large. In addition, this relationship also shows that the best suitable transformation may depend on the BOS construction. Molas and Lesaffre

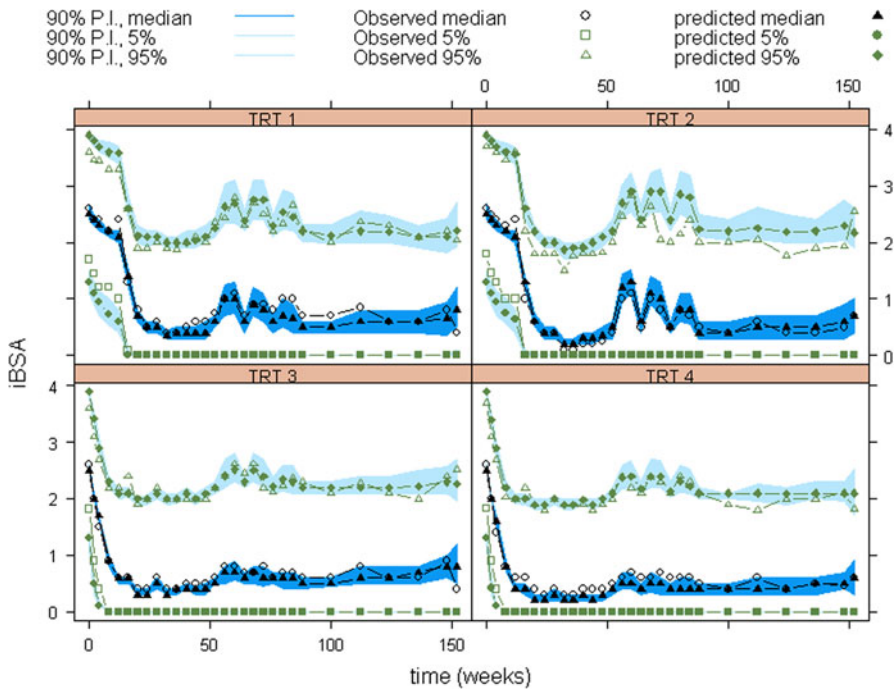


Fig. 7 VPC of the induration PK/PD model. Observed data median and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

[17] stated that when the number of possible outcomes is high the actual choice of the cut point will be less important. While this holds in the middle of the data range, it seems unlikely at the boundary where the data are J-shaped. That is, as shown in this application, transformations in link function may still be needed. Asymmetric transformations can be useful in cases when the BOS granularity is finer on one side than the other. More complex transformations than those used here could be useful if the BOS granularity is finer on some sections compared with others.

Estimating transformation parameters is known to increase computational complexity. In this application, it appears that the more computationally intensive estimation option of importance sampling in NONMEM was needed instead of the LAPLACIAN approximation to estimate two or more between-subject random effects. The computing time required was just long enough to prevent a thorough investigation of the best-fitting random effect model. It is also noted that data size in this application is relatively large, and a modest reduction could make incorporating more random effects feasible in another application. The general importance of incorporating appropriate between-subject random effects includes adjusting for within-subject correlation over time, which is important for individual simulation of data and appropriate standard error estimates. In this application, although only one

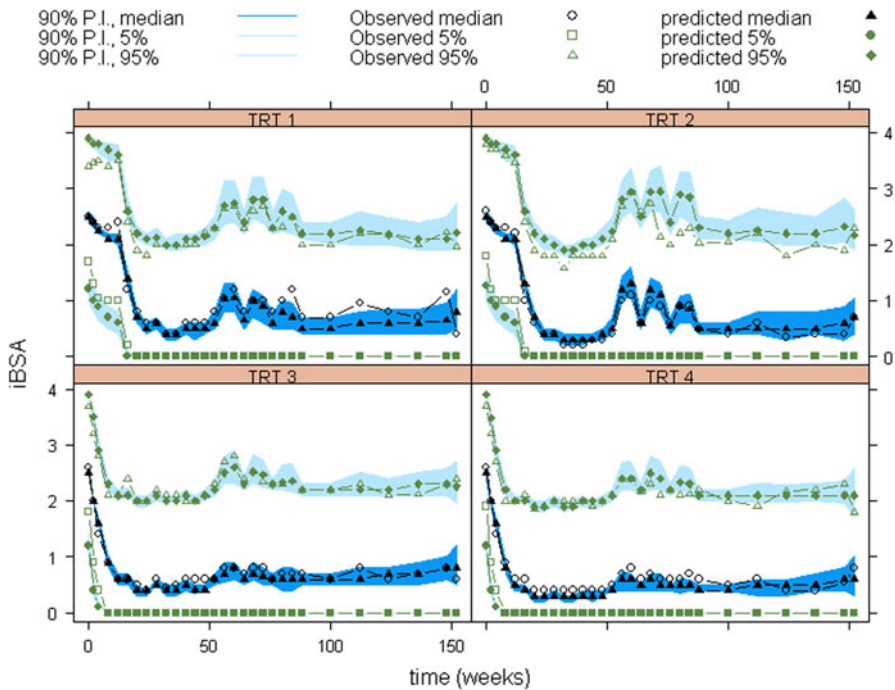


Fig. 8 VPC of the scaling pharmacokinetic/pharmacodynamic model. Observed data medians and 5th and 95th percentiles are plotted along with their corresponding predicted medians and 90% PI by treatment (TRT) groups (1 and 2 = placebo until week 12, followed by ustekinumab 45 or 90 mg; 3 = ustekinumab 45 mg; 4 = ustekinumab 90 mg)

between-subject random effect was used, the overall variability was reasonably described for all endpoints, with the exception of iBSA. The difference may be due in part to the different construction of iBSA and having a larger number of possible values, thus needing more complex transformations. This aspect does not affect the clinical objective of this analysis, but could be of interest in future investigations.

Acknowledgments The authors acknowledge Alice Zong of Centocor R&D, a division of Johnson & Johnson Pharmaceutical Research & Development, LLC. for her programming support in preparing the analysis datasets, and Dr. Rebecca E. Clemente and Mr. Robert Achenbach of Centocor Ortho Biotech Services, LLC. for their excellent assistance in preparing the manuscript. This study was funded by Centocor R&D, A Division of Johnson & Johnson Pharmaceutical Research & Development, LLC.

Appendix

The NONMEM control file of the model for iBSA is given below. Here LOSIL and LOSIU represent $z_k^{(l)}$ and $z_k^{(u)}$ in Eq. 2, and FLAG indicates whether they are smaller or greater than zero.

```

$PROB iBSA sequential, Czado
$DATA ...
$INPUT ID TIME AMT CMT DV LOSIL LOSIU FLAG EVID MDV PTIM TRT
IKA IV IK10

$$SUB ADVAN6 TOL=6
$MODEL NCOMP=3 COMP=(DEPOT) COMP=(PLASMA) COMP=(INDIR)

$PK
CALLFL = 1
  SIG = THETA(1)
  BSINCT = THETA(2) + ETA(1)
  PBMX = THETA(3)
  PBRATE = THETA(4)
  PROGRT = THETA(5)
  KOUT = THETA(6)
  EC50 = THETA(7)
  DE = THETA(8)
  PSI1 = THETA(9)
  PSI2 = THETA(10)
  A_0(3)= 1

$DES
DADT(1)= -IKA*A(1)
DADT(2)= IKA*A(1)-IK10*A(2)
  CE = A(2)/IV
  IH = CE/(CE+EC50)
DADT(3) = KOUT*(1-IH)-KOUT*A(3)

$ERROR
ZPRED = BSINCT - PBMX*(1-EXP(-PBRATE*TIME)) - DE*(1-F) + PROGRT*TIME

TY = 1
IF(FLAG.EQ.1) THEN   ; DV=0
  CZSIU = -((-LOSIU*PSI2 + 1)**(1/PSI2) - 1)
  TY = PHI((CZSIU - ZPRED)/SIG)
ENDIF
IF(FLAG.EQ.2) THEN   ; DV=1 (or 6, max)
  CZSIL = (LOSIL*PSI1 + 1)**(1/PSI1) - 1
  TY = 1 - PHI((CZSIL - ZPRED)/SIG)
ENDIF
IF(FLAG.EQ.3) THEN   ; logit(asil, asiu) <0
  CZSIL = -((-LOSIL*PSI2 + 1)**(1/PSI2) - 1)
  CZSIU = -((-LOSIU*PSI2 + 1)**(1/PSI2) - 1)
  TY = PHI((CZSIU - ZPRED)/SIG) - PHI((CZSIL - ZPRED)/SIG)
ENDIF
IF(FLAG.EQ.4) THEN   ; logit(asil, asiu) >0
  CZSIL = (LOSIL*PSI1 + 1)**(1/PSI1) - 1
  CZSIU = (LOSIU*PSI1 + 1)**(1/PSI1) - 1
  TY = PHI((CZSIU - ZPRED)/SIG) - PHI((CZSIL - ZPRED)/SIG)
ENDIF
IF(FLAG.EQ.5) THEN   ; logit(asil, asiu) include 0
  CZSIL = -((-LOSIL*PSI2 + 1)**(1/PSI2) - 1)
  CZSIU = (LOSIU*PSI1 + 1)**(1/PSI1) - 1
  TY = PHI((CZSIU - ZPRED)/SIG) - PHI((CZSIL - ZPRED)/SIG)
ENDIF
Y = -2*LOG(TY)

$THETA
  (0, 0.3) (-0.22) (0, .2) (0, 0.01) (0, 0.0002)
  (0, 0.02) (0, 0.2) (0, 0.8) (9) (4)
$OMEGA 0.1

$EST MAX=5000 NOABORT METHOD=1 LAPLACE -2LL
$COV UNCOND PRINT=E

```

References

1. Chan JR, Blumenschein W, Murphy E, Diveu C, Wiekowski M, Abbondanzo S, Lucian L, Geissler R, Brodie S, Kimball AB, Gorman DM, Smith K, de Waal Malefyt R, Kastelein RA, McClanahan TK, Bowman EP (2006) IL-23 stimulates epidermal hyperplasia via TNF and IL-20R2-dependent mechanisms with implications for psoriasis pathogenesis. *J Exp Med* 203(12):2577–2587
2. Langrish CL, Chen Y, Blumenschein WM, Mattson J, Basham B, Sedgwick JD, McClanahan T, Kastelein RA, Cua DJ (2005) IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med* 201(2):233–240
3. Yawalkar N, Karlen S, Hunger R, Brand CU, Braathen LR (1998) Expression of interleukin-12 is increased in psoriatic skin. *J Invest Dermatol* 111(6):1053–1057
4. Leonardi CL, Kimball AB, Papp KA, Yeilding N, Guzzo C, Wang Y, Li S, Dooley LT, Gordon KB (2008) Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 76-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 1). *Lancet* 371(9625):1665–1674
5. Papp KA, Langley RG, Lebwohl M, Krueger GG, Szapary P, Yeilding N, Guzzo C, Hsu MC, Wang Y, Li S, Dooley LT, Reich K (2008) Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). *Lancet* 371(9625):1675–1684
6. Hu C, Zhou H (2008) An improved approach for confirmatory phase III population pharmacokinetic analysis. *J Clin Pharmacol* 48(7):812–822
7. Zhu Y, Hu C, Lu M, Liao S, Marini JC, Yohrling J, Yeilding N, Davis HM, Zhou H (2009) Population pharmacokinetic modeling of ustekinumab, a human monoclonal antibody targeting IL-12/23p40, in patients with moderate to severe plaque psoriasis. *J Clin Pharmacol* 49(2):162–175
8. Zhou H, Hu C, Zhu Y, Lu M, Liao S, Yeilding N, Davis HM (2010) Population-based exposure-efficacy modeling of ustekinumab in patients with moderate to severe plaque psoriasis. *J Clin Pharmacol* 50(3):257–267
9. Hu C, Szapary PO, Yeilding N, Davis HM, Zhou H (2011) Informative dropout modeling of longitudinal ordered categorical data and model validation: application to exposure-response modeling of physician's global assessment score for ustekinumab in patients with psoriasis. *J Pharmacokinet Pharmacodyn* 38(2):237–260
10. Ashcroft DM, Wan Po AL, Williams HC, Griffiths CE (1999) Clinical measures of disease severity and outcome in psoriasis: a critical appraisal of their quality. *Br J Dermatol* 141(2):185–191
11. Sheiner LB, Beal SL (1985) Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. *J Pharmacokinet Biopharm* 13:185–221
12. Vonesh EF, Chinchilli VM (1997) Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York
13. Hutmacher MM, French JL, Krishnaswami S, Menon S (2011) Estimating transformations for repeated measures modeling of continuous bounded outcome data. *Stat Med* 30(9):935–949
14. Beal SL (2001) Ways to fit a PK model with some data below the quantification limit. *J Pharmacokinet Pharmacodyn* 28(5):481–504
15. Aranda-Ordaz FJ (1981) On two families of transformations to additivity for binary response data. *Biometrika* 68:357–363
16. Lesaffre E, Rizopoulos D, Tsonaka R (2007) The logistic transform for bounded outcome scores. *Biostatistics* 8(1):72–85
17. Molas M, Lesaffre E (2008) A comparison of three random effects approaches to analyze repeated bounded outcome scores with an application in a stroke revalidation study. *Stat Med* 27(30):6612–6633
18. McCullagh P, Nelder JA (1989) Generalized linear models. Chapman and Hall, London
19. Baldi I, Maule M, Bigi R, Cortigiani L, Bo S, Gregori D (2009) Some notes on parametric link functions in clinical research. *Stat Methods Med Res* 18(2):131–144
20. Box GEP, Cox DR (1964) An analysis of transformation. *J R Stat Soc B* 26:211–252
21. Czado C (1994) Parametric link modification of both tails in binary regression. *Stat Pap* 35(1):189–201
22. Hu C, Sale M (2003) A joint model for nonlinear longitudinal data with informative dropout. *J Pharmacokinet Pharmacodyn* 30(1):83–103
23. NONMEM Users Guide [Part VIII] (2008) Regents of the University of California, San Francisco

24. Zhang L, Beal SL, Sheiner LB (2003) Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *J Pharmacokinet Pharmacodyn* 30(6):387–404
25. Zou KH, Carlsson MO, Quinn SA (2010) Beta-mapping and beta-regression for changes of ordinal-rating measurements on Likert scales: a comparison of the change scores among multiple treatment groups. *Stat Med* 29(24):2486–2500