# A Novel Two-Part Mixture Model for the Incidence and Time Course of Cytokine Release Syndrome After Elranatamab Dosing in Multiple Myeloma Patients

Donald Irby[1,*], Jennifer Hibma[1], Mohamed Elmeliegy[2], Diane Wang[2], Erik Vandendries[3], Kamrine Poels[1], Blerta Shtylla[1] and Jason H. Williams[1]

**Cytokine release syndrome (CRS) is a common, acute adverse event associated with T-cell redirecting therapies such as bispecific antibodies (BsAbs). The nature of CRS events data makes it challenging to capture an unbiased exposure–response relationship with commonly used models. For example, simple logistic regression models cannot handle traditional time-varying exposure, and static exposure metrics chosen at early time points and with lower priming doses may underestimate the incidence of CRS. Therefore, more advanced modeling techniques are needed to adequately describe the time course of BsAb-induced CRS. Herein, we present a two-part mixture model that describes the population incidence and time course of CRS following various dose-priming regimens of elranatamab, a humanized BsAb that targets the B-cell maturation antigen on myeloma cells and CD3 on T cells, where the conditional time-evolution of CRS was described with a two-state (i.e., CRS-yes or no) Markov model. In the first part, increasing elranatamab exposure (maximum elranatamab concentration at first CRS event time ($C_{max,event}$)) was associated with an increased CRS incidence probability. Similarly, in the second part, increased early elranatamab exposure ($C_{max,D1}$) increased the predicted probability of CRS over time, whereas premedication including corticosteroids and IL-6 pathway inhibitors use demonstrated the opposite effect. This is the first reported application of a Markov model to describe the probability of CRS following BsAb therapy, and it successfully explained differences between different dose-priming regimens via clinically relevant covariates. This approach may be useful for the future clinical development of BsAbs.**

## Study Highlights

**WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?**

☑ Model-based optimization of dose-priming regimens for BsAbs remains a considerable challenge due to inadequate methods and an incomplete understanding of the pathophysiology of CRS.

**WHAT QUESTION DID THIS STUDY ADDRESS?**

☑ How can we effectively model the time course of CRS following BsAb monotherapy and apply this to consider or support different dose-priming regimens during drug development?

**WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?**

☑ This study highlights the applicability of the Markov model for characterizing longitudinal CRS events.

**HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?**

☑ This approach may be applied more generally for the optimization and support of dose-priming regimens for future BsAbs in clinical development.

Cancer immunotherapy has produced many promising drug modalities in the last several years, including T-cell-engaging bispecific antibodies (BsAbs).[1] These are designed to recruit and activate T cells, which involves a release of various cytokines. While BsAbs can be highly efficacious, a rapid, excessive release of cytokines can lead to unwanted side effects. Cytokine release syndrome (CRS) is a common adverse event (AE) with BsAb therapy that can present as fever and mild flu-like symptoms or more severely as hypotension and hypoxia.[2–4] The risk of CRS is highest following the first 1–2 instigating BsAb

[1]Pfizer Research and Development, Pfizer, Inc., San Diego, California, USA; [2]Oncology Research and Development, Pfizer, Inc., San Diego, California, USA; [3]Oncology Research and Development, Pfizer, Inc., Cambridge, Massachusetts, USA. *Correspondence: Donald Irby (donald.irby@pfizer.com)

Received September 20, 2024; accepted December 2, 2024. doi:10.1002/cpt.3533

doses, and then is less pronounced with subsequent dosing due to a tolerance effect. This has inspired the use of dose-priming regimens, which deliver lower initial BsAb doses to sensitize a patient's immune system while alleviating uncontrolled inflammatory responses.[5] Several BsAbs either approved for marketing or still in clinical development have adopted a dose-priming strategy to mitigate CRS.[1,5] Efforts have been made to increase the efficiency of dose selection using model-informed drug development;[6–8] however, more examples are needed that describe the dose-priming effect to better inform the study designs of future programs.

The nature of CRS events data makes it difficult to capture an unbiased exposure–response (ER) relationship with commonly used models. For example, lower priming doses that are insufficient to stimulate the immune system can delay the occurrence of CRS to times following larger step-up or maintenance doses. Simple logistic regression models cannot handle conventional time-varying exposure, and static exposure metrics chosen at early time points may underestimate the incidence of CRS. Additionally, exposure metrics calculated at the event time, such as $C_{max,event}$, can be difficult to implement, in part, because of their circular simulation logic.[9] Though longitudinal logistic regression models can implement time-varying drug concentrations, many patients will not experience CRS, making these longitudinal events data a bimodal mixture, the repeated assessments for those patients who develop CRS are likely to be autocorrelated, and it is less likely that CRS symptoms will recur after the first CRS event(s) due to a rapid drug tolerance. Therefore, more advanced modeling methodologies are needed to adequately describe the time course of BsAb-induced CRS.

Herein, we present a two-part mixture model[10–12] that describes the population incidence and time course of CRS from four phase I/II clinical trials[13] of elranatamab, a humanized BsAb that targets the B-cell maturation antigen (BCMA) on myeloma cells and CD3 on T cells, that is approved for the treatment of relapsed/refractory multiple myeloma (RRMM). A two-part mixture-modeling approach was used to incorporate the patients who did not experience CRS, where the first part of the model describes the incidence of CRS via logistic regression and the second part characterizes the conditional time-evolution of CRS with a two-state (i.e., CRS-yes or no) Markov model that also accounts for the interdependency of the events over time. Finally, the transition probabilities between the CRS-yes and -no states were adjusted with functions of time to describe response and tolerance effects. Simulations of the CRS incidence probabilities for both tested and untested dose-priming regimens were conducted to highlight the potential utility of this framework for the development of novel BsAbs.

## METHODS

### Patients and data

Longitudinal CRS data from 324 patients with RRMM were available from four phase I/II clinical trials evaluating various dosing regimens of elranatamab.[13] A brief description of these studies is provided in **Table 1**. All studies are being conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization guidelines for Good Clinical Practice. All patients provided written informed consent. The study protocols and relevant documents were approved by independent institutional review boards or ethics committees at each investigative center.

The studies included non-priming dose escalation cohorts evaluating both intravenous (IV) and subcutaneous routes of administration (0.1–1,000 μg/kg) followed by cohorts assessing single and two step-up dose priming regimens (**Table 1**). Notably, some of them differed in premedication use (dexamethasone, acetaminophen, and diphenhydramine). A daily-spaced analysis dataset was constructed from the start and end dates for each CRS event.[14] The CRS events were graded with the Common Terminology Criteria for Adverse Events v4.03 (IV cohort only) or American Society for Transplantation and Cellular Therapy systems (see **Table 1**).[4] Data were considered out to 22 days after the first elranatamab dose since most CRS events (98.5%) occurred within this timeframe.

Early and event time-based elranatamab exposure metrics (free analyte; $C_{max,D1}$ and $C_{max,event}$, respectively) obtained from a

**Table 1 Clinical studies included in the population exposure–response analyses**

| Protocol | Dosing regimen | CRS grading system | *N* | Overall CRS incidence (%)[a] | Premedication use (%)[b] | Tocilizumab/siltuximab use (%)[a] |
|---|---|---|---|---|---|---|
| C1071001 | | | | | | |
| Part 1 | IV[c]: 0.1, 0.3, 1, 3, 10, 30, and 50 μg/kg QW | CTCAE v.4.03 | 23 | 47.8 | 0.0 | 13.0 |
| | SC: 80, 130, 215, 360, 600, and 1,000 μg/kg QW | ASTCT | 30 | 73.3 | 6.7 | 30.0 |
| Part 1.1 | SC: 600 μg/kg D1 then 1,000 μg/kg QW (or Q2W) starting on D8 | ASTCT | 20 | 100.0 | 0.0 | 85.0 |
| Part 2A | SC: 44 mg D1 then 76 mg QW starting on D8 | ASTCT | 15 | 66.7 | 100.0 | 46.7 |
| C1071002 | SC: 600 μg/kg D1 then 1,000 μg/kg QW starting on D8 | ASTCT | 4 | 100.0 | 100.0 | 75.0 |
| C1071003 | | | | | | |
| Part A/B | SC: 44 mg D1 then 76 mg QW starting on D8 | ASTCT | 4 | 100.0 | 100.0 | 75.0 |
| Part A/B | SC: 12/32 mg D1/D4 then 76 mg QW starting D8 | ASTCT | 183 | 57.4 | 100.0 | 19.7 |
| C1071009 | SC: 4/20 mg D1/D4 then 76 mg QW starting D8 | ASTCT | 45 | 60.0 | 100.0 | 28.9 |

ASTCT, American Society for Transplantation and Cellular Therapy; CRS, cytokine release syndrome; CTCAE, Common Terminology Criteria for Adverse Events; D, day; N, number; IV, intravenous; QW, once weekly; SC, subcutaneous. [a]Calculated within the first 22 days of dosing. [b]A premedication cocktail (dexamethasone, acetaminophen, and diphenhydramine) administered prior to the first dose of elranatamab. [c]IV infusions administered over 1–2 hours.

semi-mechanistic target binding pharmacokinetic (PK) model[15] were considered during Markov and incidence model development, respectively. These were chosen to avoid potential bias in the estimated ER relationship while enhancing the model's simulation potential.[9,16] The median time to onset of CRS following the first dose was 2 days (interquartile range: 1–3 days). Additionally, many patients (27.5%) experienced dose delays and reductions within the first 22 days due to AEs, including CRS. Therefore, $C_{max,D1}$ was deemed to be an appropriate exposure metric for the Markov model since it is calculated prior to when the majority of the relevant events occurred and the simulated CRS event times are obtained from this model. As previously mentioned, static exposure metrics calculated at early time points may underestimate the incidence of CRS for dose-priming regimens with lower initial doses. Therefore, $C_{max,event}$ was considered for the incidence model since these exposures would be calculated using the simulated CRS event times, and many more dose-priming regimens could be simulated as a result. During model fitting, the maximum exposure on the day following the main CRS-instigating dose per cohort was used for patients without a CRS event to ensure comparable temporal scales as the long predicted half-life of elranatamab makes longer-term measures of exposure larger, by default.[9,15,17]

In addition to elranatamab exposure, other potential covariates that may influence CRS event outcomes were evaluated in the incidence and Markov models. These included baseline covariates such as soluble BCMA (sBCMA), demographic factors, prior and concomitant medications, extramedullary disease (EMD) status based on investigator assessment, and other clinically relevant factors, summarized in **Table S1**. Missing data for sBCMA, race, and EMD status were imputed given the percentage of missing values was ≤20%[18] in each case (see **Table S1**). The median and mode were used for continuous and categorical covariates, respectively. Other relevant time-varying covariates were considered in the Markov model, including tocilizumab/siltuximab use (TOCI) and the CRS outcome in the antepenultimate record (second-order Markovian element).[19] Of note, neither tocilizumab nor siltuximab was administered prophylactically to these patients; thus, any effect of their use is confounded with the CRS event. Since both tocilizumab and siltuximab have long half-lives,[20] a constant pharmacodynamic effect was explored starting a day after their initial administration.

### Software

All model building was performed in NONMEM version 7.5.0 using the FOCE method with the Laplace and Likelihood options.[21] Data pre- and post-processing were performed in R version 4.1.3.[22] The sampling importance resamplin (SIR) procedure was performed in PsN version 5.3.0.[23,24]

### Incidence model

The incidence of CRS was described using a mixture model (**Figure 1**).[10–12] Specifically, patients who experienced CRS were assigned to the subpopulation informing the Markov model parameters, and those who did not experience CRS were fixed in the CRS-no state. The probability that a patient would experience CRS was modeled via logistic regression (Eq. 1):

$$\text{logit}\left[P\left(\text{CRS}_i = 1\right)\right] = \beta_0 + \beta_n \cdot X_n \quad (1)$$

where $\text{CRS}_i$ denotes the Bernoulli random variable for patient $i$ taking the value of 1 if the patient experiences CRS and 0 otherwise, $\beta_0$ indicates the baseline logit probability for the incidence of CRS, and $\beta_n$ represents the population mean change in the log-odds of CRS for every 1 unit increase in $X_n$ (continuous) or the reference category of $X_n$ compared with its baseline characteristic (categorical). Note that Eq. 1 does not contain a random effects term as there is only one overall CRS outcome per patient.

The potential covariates outlined in **Table S1**, along with $C_{max,event}$, were tested to improve the fit of the incidence model by adding linear terms to the logit function (i.e., $\beta_n \bullet X_n$ in Eq. 1). The base model was first determined with $C_{max,event}$. Both the linear and log-transformed scales were considered. Statistical significance, goodness-of-fit, and simulation potential were used to inform this choice. The additional covariates were then added in a stepwise forward selection and backward elimination procedure. Statistical significance was determined using the likelihood ratio test for nested models, with levels $\alpha = 0.05$ and $\alpha = 0.01$ for inclusion and exclusion, respectively.

### Markov model

The conditional time-evolution of CRS was described with a two-state Markov model (see **Figure 1**). The first state represents grade 0 CRS (CRS-no; 0), and the second corresponds to grade ≥1 CRS (CRS-yes; 1). All patients start out in the CRS-no state. A discrete-time Markov model (DTMM) structure was chosen based on the frequency of the CRS assessments in the included clinical trials and the assumption of a similar data structure for predictive purposes. It has been shown that the DTMM (compared with the continuous-time Markov model, or

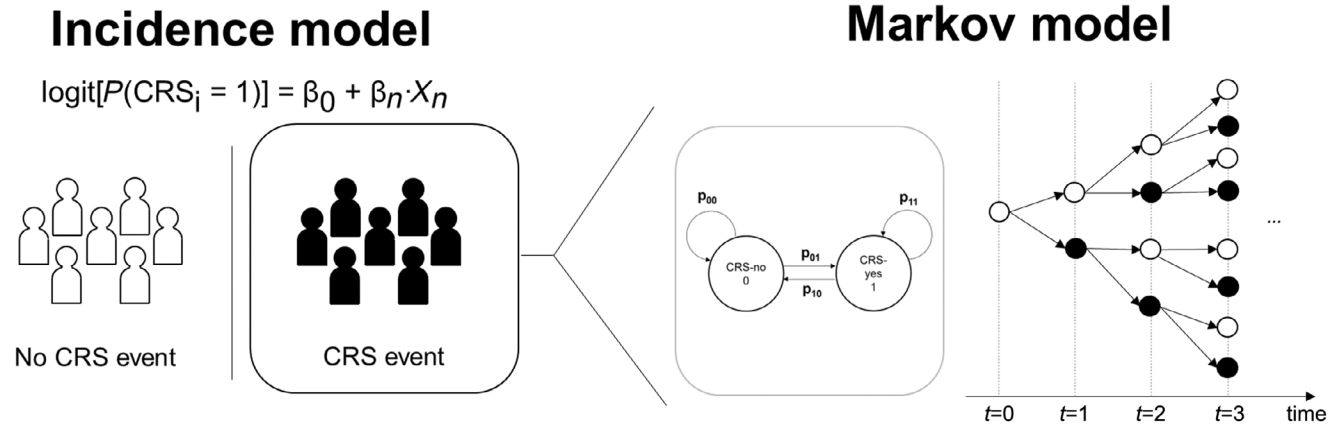


**Figure 1** Two-part mixture model of CRS. The incidence of CRS is described by a logistic regression model where the probability that a patient will have at least one CRS event is dependent upon a baseline probability (or "intercept," $\beta_0$) as well as covariates, including, but not limited to, drug exposure (or dose). In patients having at least one CRS event, a two-state Markov model (i.e., CRS-yes; 1, or no; 0) describes the evolution of CRS events over discrete-time intervals where the probability to transition between states, or to stay within a state, is estimated from the observed transitions within each patient.

CTMM) is sufficient in the cases where evenly spaced data are available.[14] The transition probabilities were modeled according to Eqs 2–7 and their corresponding terms were indexed with two digits, the first for the prior state and the second for the current state.

$$\text{logit}_{01} = \text{LGT}_{01} + f(t)_{01} \tag{2}$$

$$P_{01} = \frac{e^{\text{logit}_{01}}}{1 + e^{\text{logit}_{01}}} \tag{3}$$

$$P_{00} = 1 - P_{01} \tag{4}$$

where $\text{LGT}_{01}$ represents the baseline logit probability for the 0–1 transition and $f(t)_{01}$ the response effect starting with the first elranatamab dose.

$$\text{logit}_{10} = \text{LGT}_{10} + f(t)_{10} \tag{5}$$

$$P_{10} = \frac{e^{\text{logit}_{10}}}{1 + e^{\text{logit}_{10}}} \tag{6}$$

$$P_{11} = 1 - P_{10} \tag{7}$$

where $\text{LGT}_{10}$ represents the baseline logit probability for the 1–0 transition and $f(t)_{10}$ the tolerance effect starting after the first elranatamab dose. The response and tolerance effects after the first elranatamab dose took the form of simplified Bateman functions utilizing individual change-points in their elimination processes (Eqs 8–10).[25]

$$f(t)_{01} = \theta_{01} \cdot \left(e^{-\text{ke}_{01} \cdot \text{TAE}} - e^{-\text{ka}_{01} \cdot t}\right) \tag{8}$$

where $\theta_{01}$ represents the typical stimulation effect, $\text{ka}_{01}$ ($1/\text{MAT}_{01}$) and $\text{ke}_{01}$ ($1/\text{MET}_{01}$), the absorption and elimination rate constants for the stimulation effect, respectively, and TAE was defined according to Eq. 9:

$$\text{TAE} = \begin{cases} 0 \text{ for } t < \text{FDV} \\ t - \text{FDV for } t \geq \text{FDV} \end{cases} \tag{9}$$

where FDV corresponds to the time of first CRS event, per patient.

$$f(t)_{10} = \theta_{10} \cdot e^{-\text{ke}_{10} \cdot \text{TAE}} \tag{10}$$

where $\theta_{10}$ represents the typical desensitization effect and $\text{ke}_{10}$ ($1/\text{MET}_{10}$) the rate constant for desensitization. The value of $\theta_{10}$ was fixed to −10, making the probability of a 1–0 transition approximately 0% at the start of treatment to improve model stability and convergence while reflecting the fact that all patients start out in the CRS-no state.

Additive and exponential terms for inter-individual variability (IIV) were evaluated in the logit equations for the transition probabilities and on the Bateman function parameters, respectively. The exponential variability model was implemented as in Eq. 11:

$$P_x = \theta_x \cdot e^{\eta_j} \tag{11}$$

where $P_x$ represents one of the Bateman function parameters, $\theta_x$ is the typical value for this parameter in the population, and $\eta_j$ is the random effect around that parameter for patient j and is assumed to have a normal probability distribution with mean 0 and variance $\omega^2$.

The same covariates evaluated in the incidence model and the time-varying covariates described in the patients and data subsection were tested to improve the fit of the Markov model. They were assessed by adding linear terms to the logit equations for the transition probabilities and by modifying the parameters in the Bateman functions. The base Markov model was first determined by testing $C_{\text{max,D1}}$. Following a univariate screen, the remaining covariates were added to the model using the stepwise approach described for the incidence model. The following parameterization was used for categorical covariates modifying the Bateman function parameters (Eq. 12):

$$\theta_{\text{CatEff,x}} = \theta_{\text{cat,x}}^{\text{Cat}_x} \tag{12}$$

where $\theta_{\text{CatEff,x}}$ represents the population effect, $\theta_{\text{cat,x}}$ for the reference characteristic of $\text{Cat}_x$ compared with its baseline value. The continuous covariates modifying the Bateman function parameters were modeled according to Eq. 13:

$$\theta_{\text{ConEff,x}} = \left(\frac{\text{Con}_x}{\text{Con}_{\text{med,x}}}\right)^{\theta_{\text{con,x}}} \tag{13}$$

where $\theta_{\text{ConEff,x}}$ represents the population effect, $\theta_{\text{con,x}}$ for $\text{Con}_x$ normalized by its median value in the dataset, $\text{Con}_{\text{med,x}}$.

## Model evaluation

The adequacy of the final model was evaluated with a variety of methods. Predicted probabilities of CRS were generated over the range of observed values for $C_{\text{max,event}}$. Longitudinal typical value predictions were performed by simulating ($N = 1{,}000$) new parameter sets with the Markov model variance–covariance matrix and combining them with the exposure metrics and covariate data from the CRS-yes subset. Population-averaged joint probabilities for the proportion of CRS events and state transitions over time following the various dose-priming regimens of elranatamab were obtained by multiplying percentiles (2.5th, 50th, and 97.5th) of the predicted CRS and state transition rates by the covariate-adjusted typical value estimate for the CRS mixture probability.[10,11] Cumulative joint-predicted probabilities of CRS were obtained in a similar way. SIR[26] was used to obtain nonparametric estimates of the parameter uncertainty. Multicollinearity in the predictors was assessed by inspection of the condition number.[27]

## Simulations

Additional dosing simulations were performed according to a procedure like the one described for the longitudinal typical value predictions. Of note, and as previously mentioned, the simulated CRS event times were first obtained from the Markov model. The event time information was then used with the elranatamab PK model to calculate $C_{\text{max,event}}$. Finally, the simulated CRS incidences were obtained using the logistic regression model. A total of 200 Markov model parameter sets and 200 patients (sampling with replacement using the 324 patients) were simulated for each dosing scenario, which included all permutations of the following step-up dose 1 (sequence from 4 to 44 mg by 4 mg increments) and step-up dose 2 (sequence from 0 mg to 32 mg by 4 mg increments) proposals (99 data points total) on Days 1 and 4, respectively, with 76 mg on Day 8. All were produced with 100% premedication use.

It was observed that small deviations in the median FDV and $C_{\text{max,event}}$ resulted in large differences between the observed and predicted probabilities of CRS. This is likely due to the stochastic component of the Markov simulations and the steepness of the ER relationship in the incidence model, respectively. Therefore, it was deemed necessary to change the step size for each simulation as a function of the initial priming dose while maintaining a daily observation minimum to reduce some of its random nature. In general, one should not simulate outcomes from a DTMM with different observation frequencies than were used to develop it as the discrete Markov property would not extend to these situations.[14] However, simulating smaller steps should be acceptable in our case since we are

implementing a weaker Markovian element than should be expected with a higher observation frequency.

## RESULTS

The relevant characteristics of the 324 patients with RRMM are summarized in **Table S1**. Of note, premedication (dexamethasone, acetaminophen, and diphenhydramine) was utilized for initial elranatamab doses in 76.9% of these patients. A description of the included clinical trials is provided in **Table 1** and in a previous publication.[13] Importantly, and as previously mentioned, the data for this analysis were considered out to 22 days after the first elranatamab dose. The observed CRS incidence was 60.0%, 57.4%, and 88.4% for 4/20/76 mg, 12/32/76 mg, and 44/0/76 mg, respectively, and it was 62.7% overall. Additionally, tocilizumab/siltuximab was administered in 28.9%, 19.7%, and 69.8% of the patients undergoing 4/20/76 mg, 12/32/76 mg, and 44/0/76 mg of elranatamab, respectively.

The structure of the two-part mixture model describing the population incidence and time course of CRS following the tested regimens of elranatamab is shown in **Figure 1**. Longitudinal response and tolerance effects after the first elranatamab dose were captured by adjusting the Markov state transition probabilities with Bateman functions. Increased maximum elranatamab concentration at the first CRS event ($C_{max,event}$) and early maximum elranatamab concentration ($C_{max,D1}$) were associated with an increased CRS incidence probability and an increased probability of CRS over time, respectively. Additionally, premedication and IL-6 pathway inhibitors use both decreased the probability of CRS over time. The final model parameters were estimated with reasonable precision, and the direction of their covariate effects were in line with expectations (**Table 2**).

**Table 2 Final model parameter estimates**

| Parameter | Estimate (95% CI), covariance step | Median (95% CI), SIR |
|---|---|---|
| $LGT_{01}$ | −5.58 (−6.36, −4.80) | −5.58 (−6.32, −4.83) |
| $LGT_{10}$ | −1.58 (−1.83, −1.33) | −1.59 (−1.85, −1.32) |
| $MET_{01}$ | 4.17 (2.63, 5.70) | 4.15 (2.76, 5.76) |
| $MET_{10}$ | 0.55 (0.37, 0.73) | 0.55 (0.39, 0.73) |
| $\Theta_{01}$ | 7.40 (6.31, 8.49) | 7.38 (6.38, 8.39) |
| $\Theta_{10}$ | 10.00 (10.00, 10.00) | 10.00 (10.00, 10.00) |
| $MAT_{01}$ | 0.72 (0.55, 0.89) | 0.73 (0.57, 0.89) |
| PROB | −8.82 (−10.95, −6.68) | −8.80 (−10.86, −7.04) |
| $C_{max,event}$ on PROB | 1.82 (1.39, 2.24) | 1.82 (1.46, 2.21) |
| $C_{max,D1}$ on $\Theta_{01}$ | 0.16 (0.11, 0.22) | 0.16 (0.11, 0.22) |
| $C_{max,D1}$ on $LGT_{10}$ | 0.33 (0.13, 0.52) | 0.33 (0.16, 0.54) |
| $C_{max,D1}$ on $MET_{10}$ | 0.30 (0.07, 0.53) | 0.30 (0.08, 0.54) |
| TOCI on $LGT_{10}$ | 0.75 (0.39, 1.12) | 0.76 (0.40, 1.15) |
| TOCI on $MAT_{01}$ | 5.01 (2.19, 7.84) | 4.95 (2.64, 7.38) |
| PREMED on $MET_{10}$ | 0.70 (0.50, 0.90) | 0.70 (0.52, 0.91) |
| PREMED on $\Theta_{01}$ | 0.69 (0.60, 0.79) | 0.69 (0.60, 0.78) |

CI, confidence interval; PREMED, premedication with dexamethasone, acetaminophen, and diphenhydramine; PROB, baseline logit probability for CRS in the mixture model; SIR, sampling importance resampling; TOCI, tocilizumab/siltuximab use.

### Incidence model

Peak elranatamab exposure at FDV ($\log(C_{max,event})$) was retained in the incidence model ($C_{max,event}$ on the probability of CRS (PROB); see **Table 2**). Its odds ratio (OR; 6.2) suggests a 520% increase in the odds of CRS occurring for every 1 unit increase in $\log(C_{max,event})$ (~threefold increase in $C_{max,event}$). Of note, only 11 of the 324 patients (3.4%) dropped out during this analysis window prior to experiencing CRS; thus, any potential bias on the estimate for CRS incidence was considered negligible. The area under the ROC curve (AUROC) for the model was 0.873 (0.833–0.912 95% CI).[28] No other covariates were statistically significant.

### Markov model

Premedication with corticosteroids (PREMED), early peak elranatamab exposure ($C_{max,D1}$), and time-varying TOCI were identified as statistically significant covariates on several parameters in the Markov model (see **Table 2**). Premedication use was associated with a 30% reduction in $\theta_{01}$ and a 43% increase in $ke_{10}$ (30% decrease in $MET_{10}$), decreasing the maximum stimulation effect and expediting CRS desensitization. The impact of premedication can be visualized by comparing the with and without premedication panels for each dose-priming regimen in **Figure S2**.

Elranatamab $C_{max,D1}$ increased $\theta_{01}$ by 52% when comparing the extremes of the observed data (90th percentile: 488 ng/mL, and 10th percentile: 36 ng/mL), demonstrating an increased stimulation effect with increasing exposure. Similarly, it decreased $ke_{10}$ by 15% (119% increase in $MET_{10}$) comparing the extremes, displaying a negative association between the rate of desensitization and elranatamab exposure. Elranatamab $C_{max,D1}$ was also included in the logit equation for the 1–0 transition probability. The OR comparing the extremes of $C_{max,D1}$ (2.6) reflects a 160% increase in the odds of a 1–0 transition occurring for the increased exposure, which could reflect the increased probability of a preceding 0–1 transition. Similarly, the overall impact of elranatamab $C_{max,D1}$ can be demonstrated by comparing the 4/20/76 mg, 12/32/76 mg, and 44/0/76 mg columns of **Figures 3, 4**, and **Figure S2**.

TOCI was associated with an 80% decrease in $ka_{01}$ (501% increase in $MAT_{01}$) reducing the rate of CRS stimulation. It was also included in the logit equation for the 1–0 transition probability. Its OR (2.1) suggests a 110% increase in the odds of a 1–0 transition occurring, which is consistent with the intended use of these agents. The predicted time course of CRS for the 4/20/76 mg, 12/32/76 mg, and 44/0/76 mg dose-priming regimens shown in **Figures 3** and **4** was simulated with ~29%, 20%, and 70% chances of TOCI, respectively, to be consistent with the overall observed tocilizumab/siltuximab usage rates (**Table 1**). Again, the simulated effects of extreme tocilizumab/siltuximab usage can be referenced in **Figure S2**.

### Model diagnostics

Various additional diagnostics were used to assess the adequacy of the final model. It was stable with an acceptable degree of

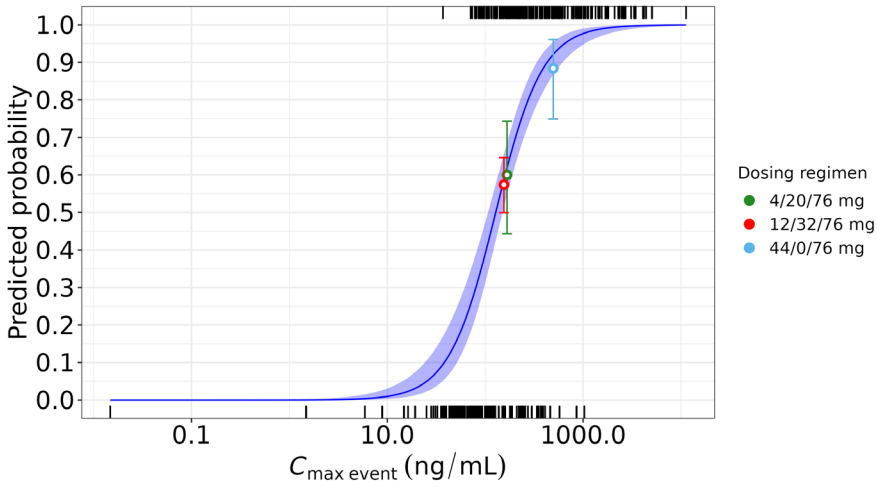


**Figure 2** Population-predicted incidence of CRS over the observed range of $C_{max,event}$. The open circles and whiskers represent the observed proportions and 95% CIs of grade ≥1 CRS at the median observed $C_{max,event}$ values for each of the dose-priming regimens. The solid blue line and ribbon represent the predicted probabilities and 95% CI of grade ≥1 CRS, respectively. The rug lines indicate the $C_{max,event}$ values for those patients who did (top) and did not (bottom) experience grade ≥1 CRS. The dosing of these amounts occurred on Days 1, 4, and 8.

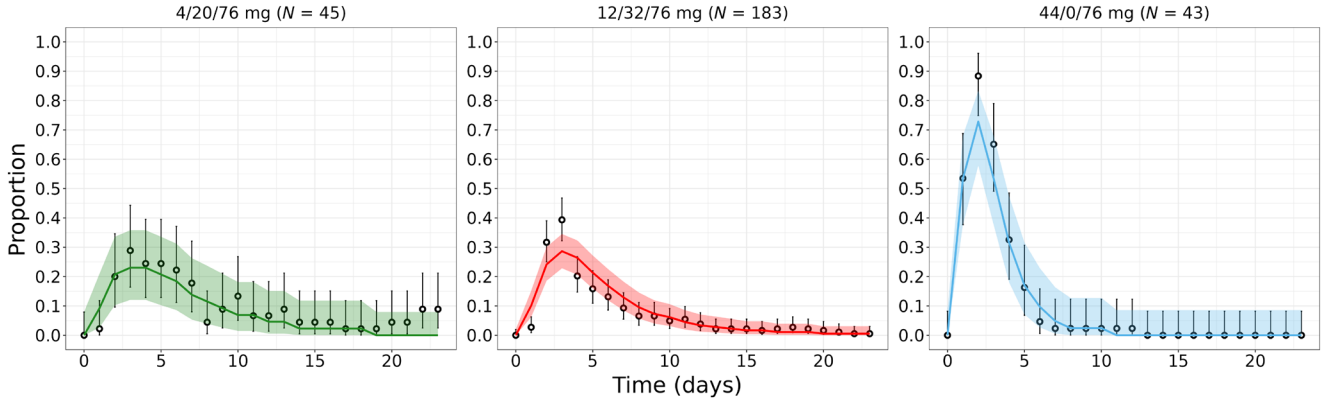


**Figure 3** Population-averaged joint-predicted CRS probabilities over time. Open circles and whiskers represent the observed fraction and 95% CI of grade ≥1 CRS, respectively. The number of patients contributing to the observed summaries per dose-priming regimen are given in the panel titles. Solid lines and shaded regions represent the median and 95% CI of the predictions of grade ≥1 CRS, respectively. The dosing of these amounts occurred on Days 1, 4, and 8.

multicollinearity in its predictors as evidenced by its condition number (<1,000).[27] The first-order Markov element adequately accounted for autocorrelation in the CRS events data with no further improvement in the objective function value with a second-order term. There were no IIV terms in the final model owing to the associated decrease in model stability and lack of strong evidence for inclusion based on the Bayesian information criterion.[29] The predicted probabilities of CRS over the observed range of elranatamab exposure captured the observed data well (**Figure 2**). Similarly, the population-averaged joint-predicted probabilities for both the proportion of CRS events and state transitions over time were consistent with the observations following different dose-priming regimens of elranatamab (**Figures 3** and **4**, respectively). Lastly, the cumulative joint-predicted probabilities of CRS also performed well in recapitulating the observed times-to-first CRS event, although with slight underprediction for the 4/20/76 mg regimen (**Figure S1**).

### Simulations

The predicted CRS incidences for a wide range of dose-priming regimens are presented as a three-dimensional surface in **Figure 5**. The CRS probability surface includes a region where two-step regimens greatly reduced the predicted probability of CRS compared with most one-step (i.e., X/0/76 mg) regimens. Additionally, the simulations indicated that a first step-up dose of at least 12 mg resulted in a minimal dependence of CRS on doses 2 and 3, although initial doses greater than 12 mg increased the probability of CRS. Conversely, for dose 1 amounts less than 12 mg, the overall incidence increased with increasing amounts of dose 2, yet never exceeded the predicted probability of 12 mg on dose 1.

## DISCUSSION

In the past 3 years, the United States Food and Drug Administration has approved eight BsAbs for oncology indications, highlighting the increasing role of these modalities in

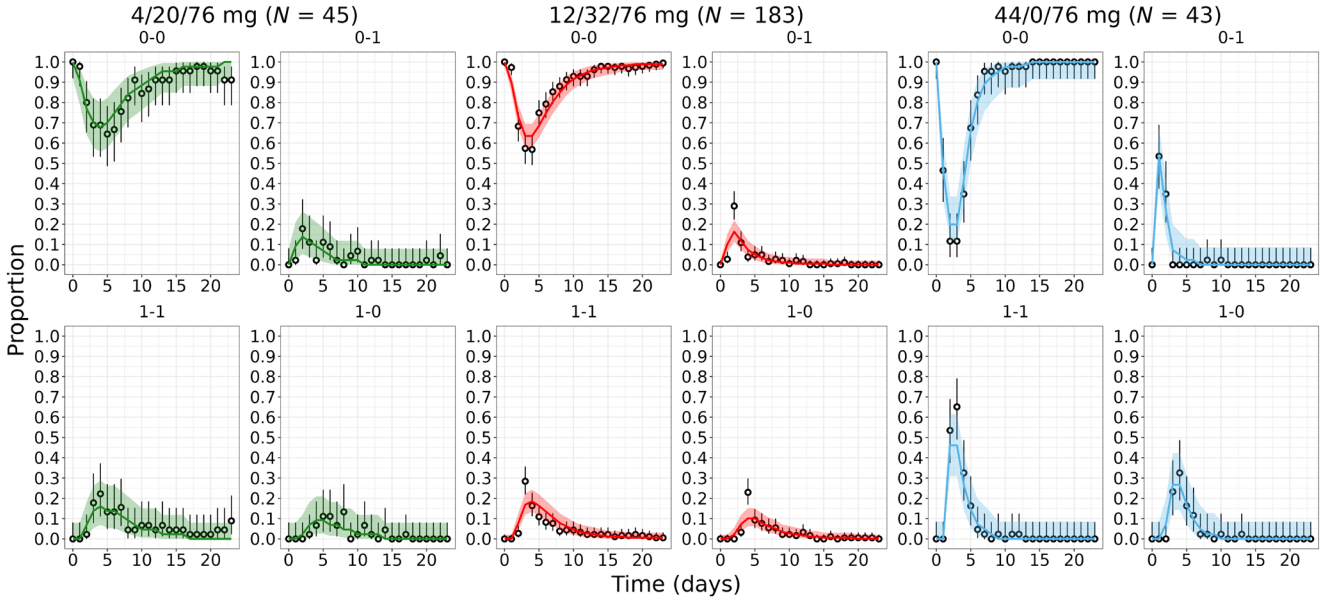


**Figure 4** Population-averaged joint-predicted CRS transition probabilities over time. Open circles and whiskers represent the observed fraction and 95% CI of each transition possibility for grade ≥1 CRS, respectively. The number of patients contributing to the observed summaries per dose-priming regimen are given in the panel titles. Solid lines and shaded regions represent the median and 95% CI of the predictions for each transition of grade ≥1 CRS, respectively. The dosing of these amounts occurred on Days 1, 4, and 8.

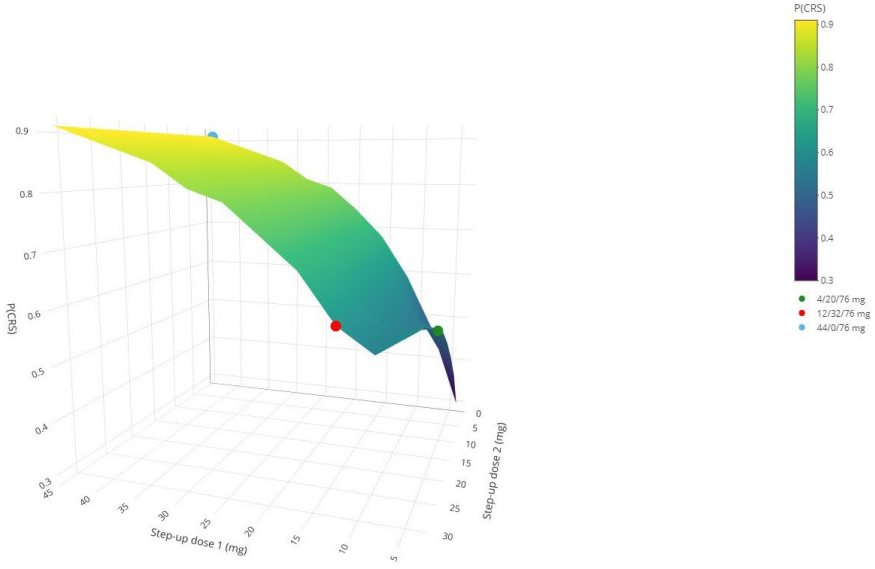


**Figure 5** Population-predicted CRS incidence over a wide range of dose-priming regimens. The CRS incidence probabilities were generated using all permutations of the following step-up dose 1 (sequence from 4 to 44 mg by 4 mg) and step-up dose 2 (sequence from 0 to 32 mg by 4 mg) proposals (99 data points total) on Days 1 and 4, respectively, with 76 mg on Day 8. All were produced with 100% premedication usage. The observed dose-priming regimens (4/20/76, 12/32/76, and 44/0/76 mg) are plotted as green, red, and blue dots, respectively.

cancer treatment.[30] While the treatments are promising, CRS can lead to life-threatening complications that often require hospitalization and can limit access to these therapies in community settings. Model-based approaches that adequately describe the dose-priming effect could be useful to simulate the expected CRS profiles for proposed treatment regimens.

Several aspects of CRS events data following novel BsAb treatment regimens make it difficult to adequately characterize using traditional approaches, that is, logistic regression. Longitudinal models are needed to address these issues, although these come with challenges of their own. For example, many patients do not experience CRS, and their inclusion in a longitudinal model can bias its parameter estimates, once again resulting in an underprediction of CRS.[12] However, two-part mixture models can alleviate this problem by separating the CRS-yes and -no patients between different parameter sets.[10,11] Additionally, the CRS events data for

a given patient will be autocorrelated (current observation is dependent on the previous one(s)), and implementing a model that assumes they are independent will result in undesirable simulation properties (i.e., spurious CRS events at random times and shorter-than-expected durations). Such influence between successive observations, given they are evenly spaced in time, can be handled with a first-order DTMM.[19] Lastly, in general, it is less likely that CRS symptoms will recur (following their resolution) after the first CRS event(s) due to a rapid drug tolerance.[3] This early increased, then decreased risk of CRS must be accounted for and may be characterized with simplified Bateman functions.[25] Our two-part mixture model was successful in overcoming these challenges and in reproducing various features of the observed CRS data.

Several covariate effects were identified in the Markov model, and the direction of their parameter estimates were as expected (**Table 2** and **Figure S2**). First, increased initial elranatamab exposure was predicted to increase the stimulation effect for CRS ($\theta_{01}$). Premedication with corticosteroids was found to decrease the stimulation effect and expedite the rate of desensitization ($ke_{10}$), which supports the implementation of premedication regimens with initial doses as well as the use of their agents in management of these CRS events. Lastly, TOCI was predicted to slow the rate of absorption ($ka_{01}$) of the stimulation effect and increase the probability of a 1–0 transition, reducing the probability of CRS over time. Similarly, this reinforces the use of IL-6 pathway inhibitors in the clinic to manage BsAb-induced CRS and indicates a potential utility in using these treatments prophylactically.

The simulated CRS response surface (**Figure 5**) highlighted the dependence of CRS on dose 2 when the initial dose was low (4–8 mg) yet was independent of dose 2 when the initial dose was greater than or equal to 12 mg, providing additional perspective and support for the approved elranatamab step-up dose regimen (12/32/76 mg). This indicates that the selected two step-up priming regimen of 12/32/76 mg provides the lowest initial dose that eliminates the dependency on doses 2 or higher. This supports the predictable profile of this regimen and justifies the recommended hospitalization guidelines for elranatamab.

Further, while the surface plot may suggest a possible reduction in the probability of any grade CRS for some untested dose-priming regimens, it is currently uncertain whether such theoretical regimens would result in an undesirable increase in the proportion of grade ≥2 CRS. A 3-state DTMM capable of distinguishing grades 0, 1, and 2/3 CRS would therefore be highly valuable as reducing the risk of grade 2/3 CRS is a critical safety benchmark for developing BsAbs. This kind of model was considered but not implemented due to the low number of observed transitions into or out of the grade 2/3 state when compared with the total number of transitions (<3%, see **Table S2**). Future work may include fixing the transition probabilities for the grade 2/3 state to near-final estimates[14] or to their observed values,[31] or to obtain and include more data in the analysis. Alternatively, a minimal CTMM could be explored.[32]

While our model was successful in describing CRS following different dose-priming regimens of elranatamab, there are still some limitations to its use. First, although the incidence model has a "good" AUROC (0.873),[28] its discrimination could be improved to facilitate individual-level dosing recommendations. Further advancements on this front may be accomplished with additional data or alternative modeling methods. Specifically, we were unable to adequately test the potential covariate effects of various cytokine measures in the models. Should early changes in cytokine levels following the initial elranatamab doses improve the discrimination of the incidence model, for example, then follow-up work may include synergizing with a mechanistic model for BsAb-induced cytokine release. Additionally, the dosing simulations will not be as reliable outside the observed range of step-up dose 1 (4–44 mg) since the Bateman functions involved in the generation of FDV are modified using $C_{max,D1}$. Other noteworthy modeling considerations are included in the supplement.

To the best of our knowledge, this is the first reported application of a Markov model to describe the time course of CRS following BsAb monotherapy. It successfully explained differences in the time course of CRS between the various dose-priming regimens of elranatamab via clinically relevant covariates. This modeling approach demonstrates how the combination of PK, CRS, and other covariate data, integrated within the mixture-modeling framework, could be utilized to support future BsAb programs. That is, it may be possible to shrink the dose optimization space and expedite drug development by exploring the CRS-related properties of additional dose-priming regimens *in silico*.

## SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).


## ACKNOWLEDGMENTS

The authors thank the MagnetisMM trial patients and their families, as well as the study investigators, nurses, and site staff. This study was sponsored by Pfizer.

## FUNDING

Study was funded by Pfizer, Inc.


## CONFLICT OF INTEREST

D.I., J.H., M.E., D.W., E.V., K.P., and B.S. are full-time employees of Pfizer, Inc., and own Pfizer stock. J.W. was a full-time employee of Pfizer, Inc. at the time of conducting this analysis and may own Pfizer stock.

## AUTHOR CONTRIBUTIONS

D.I. performed the research. D.I., J.H., and J.W. wrote the manuscript. D.I., J.H., and J.W. designed the research. D.I., J.H., M.E., D.W., E.V., K.P., B.S., and J.W. analyzed the data.

## DATA AVAILABILITY STATEMENT

Upon request and subject to review, Pfizer will provide the data that support the findings of this study. Subject to certain criteria, conditions, and exceptions, Pfizer may also provide access to the related individual de-identified participant data. See https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information.

1. Carrasco-Padilla, C., Hernaiz-Esteban, A., Alvarez-Vallina, L., Aguilar-Sopena, O. & Roda-Navarro, P. Bispecific antibody format and the Organization of Immunological Synapses in T cell-redirecting strategies for cancer immunotherapy. *Pharmaceutics* **15**, 132 (2022).
2. Li, J. *et al*. CD3 bispecific antibody-induced cytokine release is dispensable for cytotoxic T cell activity. *Sci. Transl. Med.* **11**, eaax8861 (2019).
3. Shimabukuro-Vornhagen, A. *et al*. Cytokine release syndrome. *J. Immunother. Cancer* **6**, 56 (2018).
4. Lee, D.W. *et al*. ASTCT consensus grading for cytokine release syndrome and neurologic toxicity associated with immune effector cells. *Biol. Blood Marrow Transplant.* **25**, 625–638 (2019).
5. Chen, X., Kamperschroer, C., Wong, G. & Xuan, D. A modeling framework to characterize cytokine release upon T-cell-engaging bispecific antibody treatment: methodology and opportunities. *Clin. Transl. Sci.* **12**, 600–608 (2019).
6. Hutchings, M. *et al*. Dose escalation of subcutaneous epcoritamab in patients with relapsed or refractory B-cell non-Hodgkin lymphoma: an open-label, phase 1/2 study. *Lancet* **398**, 1157–1169 (2021).
7. Li, T. *et al*. Semimechanistic physiologically-based pharmacokinetic/pharmacodynamic model informing Epcoritamab dose selection for patients with B-cell lymphomas. *Clin. Pharmacol. Ther.* **112**, 1108–1119 (2022).
8. Li, T. *et al*. Simplifying selection and optimization of stepup dosing of subcutaneous Epcoritamab to mitigate CRS risk using repeated time-to-event modeling. *Cancer Res.* **83**, 2798 (2023).
9. Wiens, M.R., French, J.L. & Rogers, J.A. Confounded exposure metrics. *CPT Pharmacometrics Syst. Pharmacol.* **13**, 187–191 (2024).
10. Kowalski, K.G., McFadyen, L., Hutmacher, M.M., Frame, B. & Miller, R. A two-part mixture model for longitudinal adverse event severity data. *J. Pharmacokinet. Pharmacodyn.* **30**, 315–336 (2003).
11. Lavielle, M., Mesa, H., Chatel, K. & Vermeulen, A. Mixture models and model mixtures with MONOLIX. Population Approach Group Europe 19th Annual Conference <www.page-meeting.org/?abstract=1853> (2010).
12. Upton, R.N. & Mould, D.R. Basic concepts in population modeling, simulation, and model-based drug development: part 3-introduction to pharmacodynamic modeling methods. *CPT Pharmacometrics Syst. Pharmacol.* **3**, 1–16 (2014).
13. Elmeliegy, M. *et al*. Dose optimization to mitigate the risk of CRS with Elranatamab in multiple myeloma. *Blood* **140**, 7174–7175 (2022).
14. Lu, T., Yang, Y., Jin, J.Y. & Kagedal, M. Analysis of longitudinal-ordered categorical data for muscle spasm adverse event of Vismodegib: comparison between different Pharmacometric models. *CPT Pharmacometrics Syst. Pharmacol.* **9**, 96–105 (2020).
15. Hibma, J.E.E.M., King, L.E., McCush, F., Wang, D. & Williams, J.H. Population pharmacokinetic analysis of elranatamab in patients with multiple myeloma. Population Approach Group Europe 31st Annual Meeting <www.page-meeting.org/?abstract=10648> (2023).
16. Ruiz-Garcia, A. *et al*. A comprehensive regulatory and industry review of modeling and simulation practices in oncology clinical drug development. *J. Pharmacokinet. Pharmacodyn.* **50**, 147–172 (2023).
17. *Elrexfio [Package Insert]* (Pfizer, Inc., New York, NY, USA, 2023).
18. Irby, D.J. *et al*. Approaches to handling missing or "problematic" pharmacology data: pharmacokinetics. *CPT Pharmacometrics Syst. Pharmacol.* **10**, 291–308 (2021).
19. Lacroix, B.D., Lovern, M.R., Stockis, A., Sargentini-Maier, M.L., Karlsson, M.O. & Friberg, L.E. 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 (2009).
20. Leung, E. *et al*. Pharmacokinetic/pharmacodynamic considerations of alternate dosing strategies of tocilizumab in COVID-19. *Clin. Pharmacokinet.* **61**, 155–165 (2022).
21. Beal, S.L., Sheiner, L.B., Boeckmann, A.J. & Bauer, R.J. *NONMEM 7.5 Users Guides* (ICON Plc, Gaithersburg, MD, USA, 1989–2020) <https://nonmem.iconplc.com/nonmem750>.
22. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2022) <https://www.R-project.org/>.
23. Lindbom, L., Pihlgren, P. & Jonsson, E.N. PsN-toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput. Methods Programs Biomed.* **79**, 241–257 (2005).
24. Lindbom, L., Ribbing, J. & Jonsson, E.N. Perl-speaks-NONMEM (PsN)—a Perl module for NONMEM related programming. *Comput. Methods Programs Biomed.* **75**, 85–94 (2004).
25. Garrett, E.R. The Bateman function revisited: a critical reevaluation of the quantitative expressions to characterize concentrations in the one compartment body model as a function of time with first-order invasion and first-order elimination. *J. Pharmacokinet. Biopharm.* **22**, 103–128 (1994).
26. Dosne, A.G., Bergstrand, M. & Karlsson, M.O. An automated sampling importance resampling procedure for estimating parameter uncertainty. *J. Pharmacokinet. Pharmacodyn.* **44**, 509–520 (2017).
27. Bonate, P. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation* (Springer, New York, NY, USA, 2006).
28. Nahm, F.S. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J. Anesthesiol.* **75**, 25–36 (2022).
29. Mould, D.R. & Upton, R.N. Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e38 (2013).
30. Lim, K., Zhu, X.S., Zhou, D., Ren, S. & Phipps, A. Clinical pharmacology strategies for bispecific antibody development: learnings from FDA-approved bispecific antibodies in oncology. *Clin. Pharmacol. Ther.* **116**, 315–327 (2024).
31. Brekkan, A., Lledo-Garcia, R., Lacroix, B., Jönsson, S., Karlsson, M.O. & Plan, E.L. Characterization of anti-drug antibody dynamics using a bivariate mixed hidden-markov model by nonlinear-mixed effects approach. *J. Pharmacokinet. Pharmacodyn.* **51**, 65–75 (2024).
32. Schindler, E. & Karlsson, M.O. A minimal continuous-time Markov pharmacometric model. *AAPS J.* **19**, 1424–1435 (2017).