Received: 1 April 2022 | Revised: 9 August 2022 | Accepted: 11 August 2022

DOI: 10.1002/psp4.12863

ARTICLE

Dose/exposure–response modeling in dose titration trials: Overcoming the titration paradox

Niels Rode Kristensen | **Henrik Agersø**

Pharmacometrics, Novo Nordisk A/S, Søborg, Denmark

Correspondence
Niels Rode Kristensen, Novo Nordisk A/S, Vandtårnsvej 114, DK-2860 Søborg, Denmark.
Email: nikr@novonordisk.com

Funding information
Novo Nordisk

Abstract

Response-based dose individualization or dose titration is a powerful approach to achieve precision dosing. Yet, titration as an individualization strategy is underused in drug development and therefore not reflected in labeling, possibly partly because of the data analysis challenges associated with assessing dose/exposure–response under dose titration, where there is an inherent risk of selection bias because poor responders would get high doses, whereas good responders would get low doses. In a recent article, this issue of selection bias was termed the "titration paradox." In this study, we demonstrate by means of simulation that the titration paradox may be overcome if longitudinal data from dose titration trials is analyzed using a population approach that accounts for the fact that dose/exposure–response relationships differ between individuals. We show that with an appropriate sample size and missing data missing at random, stepwise dose/exposure–response modeling based on data obtained under dose titration is not by definition subject to model selection bias or bias in parameter estimates. We also illustrate the challenges of graphical exploration of data obtained under dose titration and discuss the use of model diagnostic tools with such data. Our study shows that if, at every timepoint in the course of a trial, there is a clear causal relationship between the response and the dose/exposure level, and a population approach is used, it will in many cases be possible to develop, estimate, and appropriately qualify a dose/exposure–response model also for data obtained under dose titration, thus overcoming the titration paradox.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Dose titration is underused in drug development and therefore not reflected in labeling, which may partly be because of the data analysis challenges associated with assessing dose/exposure–response under dose titration, where there is a risk of selection bias—the "titration paradox."

WHAT QUESTION DID THIS STUDY ADDRESS?

How should dose/exposure–response data from dose titration trials be analyzed to overcome the titration paradox?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

This study shows that the use of dose titration does not by definition cause issues with selection bias. If there is a clear causal relationship between dose/exposure and response, using a population approach will often allow appropriate inferences to be made.

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

If analyzed properly, longitudinal data from dose titration trials can be just as, if not more, powerful than data from fixed dose trials for the development of dose/exposure–response models to support drug development and inform labeling.

INTRODUCTION

Precision dosing, that is, individualization of drug doses to optimize the benefit–risk balance for individual patients, is receiving increasing attention. Several authors have pointed out the importance of moving away from a one-size-fits-all approach for dose selection in drug development[1] and for providing dosing recommendations in drug labeling[2] not only for the benefit of the patients but also to address the increasing focus from prescribers and payers on using the most cost-effective dose for each individual patient in the broader real-world population.[3] Precision dosing moves beyond the common adjustment of the dose based on body size, demographic factors, renal or hepatic impairment, concomitant medication, and so on and can be guided by observed drug exposure, biomarkers of response, or even observed response.[1–3]

Response-based dose individualization is particularly powerful, and simulations have shown that the responder rate, which is often modest with a single fixed dose level, can be substantially increased by individual dose titration.[4] Yet, titration as an individualization strategy is underused in pivotal trials and therefore not reflected in labeling.[2,5] An analysis of drugs approved by the US Food and Drug Administration from 2013 to 2017 revealed that only 39% of the drugs considered amenable for titration had information about this in labeling, and only for 53% of these drugs had titration been studied in pivotal trials.[5] One reason for this may be additional trial complexity not only in confirmatory trials, where more visits and longer trial durations may be needed, but also in exploratory trials, where a wider dose range and broader population may have to be studied to generate the data and develop the pharmacokinetic–pharmacodynamic and/or dose/exposure–response models needed to determine an appropriate dose titration strategy for confirmatory trials.[1] Another possible reason and an important factor to consider in general for dose/exposure–response modeling using data obtained under dose titration is how to avoid the selection bias that is inherent to such situations.

Selection bias in this context is bias in the estimated dose/exposure–response relationship due to individuals with poor response receiving high doses and individuals with good response receiving low doses. In a recent article, the issue was called "the titration paradox" because of the paradoxical results observed in an analysis of clinical data with titration of two anesthetics and one vasoactive drug, where the dose–response relationships based on average data appeared to be inverted, showing decreased response with increasing dose.[6] In addition to reporting the analysis of the clinical data, the authors show by mathematical proof and Monte Carlo simulations that a negative or zero correlation between dose and response is in fact the expected finding for the analysis of dose–response data under dose titration, when data are naïvely pooled across individuals, and they speculate that using a mixed-effects model for the analysis of multiple data points per individual could be more appropriate. Regulatory guidelines also recognize the risk of obtaining misleading or inverted dose/exposure–response curves if variability between individuals is not properly accounted for, but do not provide specific guidance for performing such analysis.[7,8] A recent tutorial on exposure–response analysis does not cover this either, as it primarily focuses on an analysis of single timepoint change-from-baseline values in parallel-group dose–response trials.[9]

Analysis of dose–response relationships based on longitudinal data from dose titration trials has previously been investigated using marginal structural models (MSM) with inverse probability of treatment weighting (IPTW)[10,11] and with dynamic linear mixed-effects models (DLME).[12] Both approaches focus on a situation where the current response depends not only on the current dose but also on previous doses trough a dependence on previous responses. Obtaining unbiased results with the MSM approach is challenging,[11] but it may be an appropriate choice in observational studies where the dose-modification algorithm is not well defined[12] because it implicitly includes the development of a model for the dose-modification process. However, when a deterministic

dose-modification algorithm is implemented, such as in well-controlled clinical trials, the DLME approach may outperform the MSM approach.[12] In formal statistical terms, the dose-modification process is independent of the measurement process in such cases, making it unnecessary to jointly model the two processes and allowing the dose to be handled as a time-varying covariate in the analysis of the observed responses when using maximum likelihood estimation.[13] This is analogous to the problem of longitudinal data analysis with missing responses, where unbiased estimates can be obtained with maximum likelihood under the missing-at-random assumption without explicitly modeling the missingness mechanism.[13]

A special case of the dose titration scenarios previously considered is when the current response only depends on the current dose,[13] and there is a clear causal relationship between the current dose and the current response, for example, a situation where the response is measured at a timepoint where steady state has been achieved for the associated effects of a given dose, whereafter the dose is titrated based on the observed response using a deterministic titration algorithm. This is, for example, the case for titration of insulin doses based on measurements of plasma glucose, where weekly adjustments of the dose of basal insulin according to a prespecified titration algorithm aiming to achieve a certain fasting plasma glucose (FPG) target has been the norm for clinical trials ever since the so-called treat-to-target trial[14] was reported. In this trial, basal insulin treatment was initiated in patients with Type 2 diabetes and inadequate glycemic control using a common starting dose for all patients followed by individual weekly dose adjustments based on patient-performed measurements of FPG using a progressive titration algorithm with an FPG target of 100 mg/dl (5.6 mmol/L), and most patients quickly and safely achieved hemoglobin A1C ≤7%, as recommended in guidelines for diabetes management.[15,16] Today, 2 decades later, the approved labels for the two most commonly used basal insulin analogs for once-daily administration (insulin glargine[17] and insulin degludec[18]) still only recommend a starting dose level and that the dose should subsequently be adjusted based on the individual's metabolic needs, blood glucose monitoring results, and glycemic control goal, but no specific dose titration algorithm is given, and, to our knowledge, dose–response models for FPG based on data from the many treat-to-target clinical trials performed for these insulin analogs have not been published. This may in part be attributed to a concern about the associated risk of selection bias and the methodological challenges with the analysis of data from dose titration trials.[6,10–12]

In this article, we demonstrate by means of a simulation study that dose–response model development based on data from a dose titration trial may be feasible without selection bias if the data are analyzed using a population approach. We use an example that mimics an exploratory trial for a basal insulin analog. We investigate model development in terms of model discrimination aspects (linear vs. nonlinear dose–response and absence vs. presence of a time effect due to disease progression) as well as bias and variance of the estimated model parameters, and we do this for two different trial designs and across different missing data scenarios. We also illustrate the challenges with graphical exploration of such data, investigate the performance of some common model diagnostic tools, and discuss how to appropriately apply these in dose titration trials.

METHODS

Simulation scenarios

The simulation example used in this article mimicked a single arm in an exploratory basal insulin trial, which are typically of 16–26 weeks duration, include 50–100 subjects, and apply weekly dose titration based on prebreakfast self-measured plasma glucose (SMPG) values measured by the subjects at home to reach a certain SMPG target (e.g., 3.9–6.0 mmol/L). The following two scenarios were simulated:

1. A rich-data scenario with 100 subjects for 26 weeks, and
2. A limited-data scenario with 50 subjects for 16 weeks.

Titration algorithm

A treat-to-target approach to dose titration was simulated with a starting dose of 10 units of insulin per day and weekly dose adjustments based on prebreakfast SMPG according to a progressive algorithm with an increase of four insulin units per day for SMPG >7.0 mmol/L, an increase of two insulin units per day for SMPG between 6.0 and 7.0 mmol/L, no change for SMPG between 3.9 and 6.0 mmol/L, and a decrease of two insulin units per day for SMPG in the hypoglycemic range <3.9 mmol/L.

Simulation models

Four different variants (with or without nonlinearity and with or without disease progression) of the following basic dose–response model were used for generating simulated data sets:

$$\mathrm{SMPG} = a - b * \frac{\mathrm{DOSE}}{1 + c * \mathrm{DOSE}} + d * \mathrm{WEEK} + \varepsilon, \quad \varepsilon \sim N(0, \Sigma)$$

The model describes the relationship between the insulin dose in a given week and the resulting SMPG. Parameter a represents the SMPG without insulin treatment, parameter b represents the glucose-lowering effect of one unit of insulin in the linear variants of the model ($c = 0$) and is equivalent to the maximum effect (E_{max}) divided by the dose producing 50% of the maximum effect (ED_{50}) in the classic E_{max} model in the nonlinear variants of the model ($c > 0$), parameter c essentially describes the degree of deviation from a linear dose–response relationship and is for the nonlinear variants of the model equivalent to $1/ED_{50}$, parameter d describes disease progression over time,[19] and variable ε represents the day-to-day variability for SMPG measurement.[20] Although a linear dose–response model has some unfortunate properties when extrapolated to high doses, it is often an appropriate model for basal insulins in the clinically relevant dose range.[21,22] The model was used in a mixed-effects version with log-normally distributed between-subject variability included on parameters a and b with a positive correlation. Although also allowing ED_{50} to differ between subjects in an E_{max} model would usually be relevant, for simplicity we decided not to include additional variability terms in the nonlinear variants of the model. The parameter values used for simulation with the four different variants of the model are summarized in Table 1. The parameter values were selected to match average observed SMPG and dose data from a clinical trial with insulin degludec.

Missing data scenarios

For each of the overall scenarios, the following five different scenarios for missing data attributed to dropout were simulated: (1) no dropout, (2) 15% random dropout, (3) 30% random dropout, (4) 15% nonrandom dropout, and (5) 30% nonrandom dropout. Random dropout was simulated by assuming that 15% or 30% of all subjects dropped out at random times after Week 2, corresponding to a missing-completely-at-random scenario in statistical literature. Nonrandom dropout was simulated by assuming that the 15% or 30% least responsive subjects (highest ratio between parameters a and b in the simulation model) dropped out at random times after Week 2, corresponding to a missing-not-at-random scenario in statistical literature (because dropout depends on the unobserved values of a and b), although it could be argued that the scenario is close to a missing-at-random scenario (if dropout had been based on observed values of a and b, e.g., preliminary estimates of the parameters).

Simulation and estimation

For each missing data scenario in each overall scenario, $N = 1000$ replicates of the trial were simulated using each of the four variants of the simulation model. Each replicate was used for estimation with the true model to evaluate estimation performance and for estimation with alternative models to evaluate model discrimination. Figure 1 shows which alternative models were tested for each true model. All tests were likelihood ratio tests based on an assumption of a χ^2 distribution with one degree of freedom, and model discrimination was evaluated by calculating empirical Type 1 (false positive) and Type 2 (false negative) error rates based on tests of each alternative model versus the true model using a p value of 0.05. From the Type 2 error rate (β),

TABLE 1 Parameter values used for simulation with the four different variants of the simulation model

Parameter	Linear model	Nonlinear model	Linear model with disease progression	Nonlinear model with disease progression
a [mmol/L][a]	10.0	10.7	10.0	10.7
b [(mmol/L)/unit][a]	0.103	0.200	0.103	0.200
c (1/unit)	0	0.015	0	0.015
d [(mmol/L)/week]	0	0	0.077	0.077
BSV – a (%CV)	20	20	20	20
BSV – b (%CV)	60	60	60	60
BSV correlation – a–b	0.5	0.5	0.5	0.5
Σ [(mmol/L)^2]	1	1	1	1

Parameter a represents the SMPG without insulin treatment, parameter b represents the glucose-lowering effect of one unit of insulin in the linear variants of the model ($c = 0$) and is equivalent to Emax/ED50 in the classic Emax model in the nonlinear variants of the model ($c > 0$), parameter c describes the degree of deviation from a linear dose–response relationship and is for the nonlinear variants of the model equivalent to 1/ED50, parameter d describes disease progression over time.Abbreviations: BSV, between-subject variability; %CV, percentage coefficient of variation.

[a]Rounded to three significant digits.

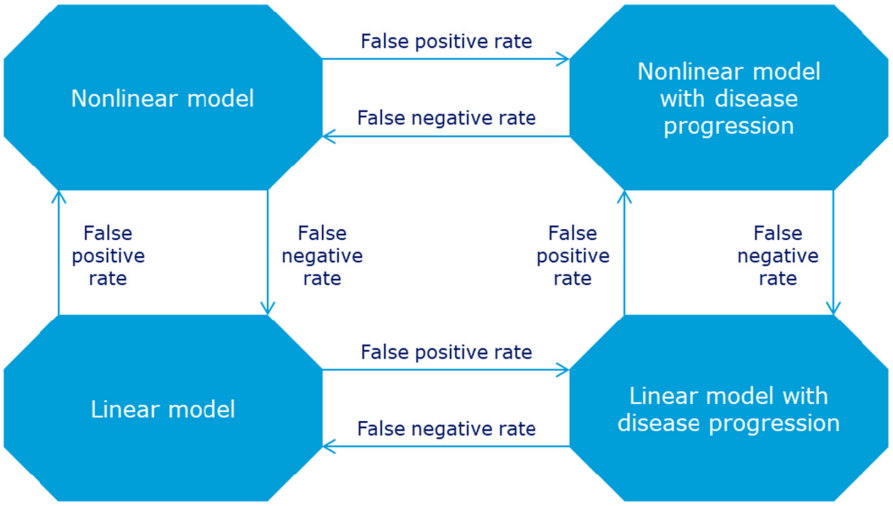

FIGURE 1 Overview of models used for simulation and estimation. Arrows indicate tests performed to investigate model discrimination performance and calculate Type 1 and Type 2 error rates.

the empirical model discrimination power was derived as $1-\beta$. Estimation performance was evaluated by calculating the relative bias for all parameters and the relative root mean square error (RMSE) and coverage probability for the fixed-effect parameters as follows:

$$\text{Relative bias}\,(\%) = 100\,\% \bullet \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\theta}_i - \theta^*}{\theta^*}$$

$$\text{Relative RMSE}\,(\%) = 100\,\% \bullet \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{\left(\hat{\theta}_i - \theta^*\right)^2}{\theta^{*2}}}$$

$$\text{Coverage probability}\,(\%) = 100\% \bullet \frac{1}{N}\sum_{i=1}^{N} I\left(\left|\hat{\theta}_i - \theta^*\right| < 1.96 \bullet \hat{\sigma}_i\right)$$

Here, $\hat{\theta}_i$ is the estimate and $\hat{\sigma}_i$ the standard deviation, θ^* is the true value, and $I()$ is a binary indicator function. Model estimation was performed in NONMEM Version 7.3 (ICON plc) using the first-order conditional estimation algorithm, and simulation, plotting, and calculation of summary statistics was performed in R Version 3.5.3 (R Foundation for Statistical Computing).

Graphical data exploration and model diagnostics

To illustrate the challenges with graphical analysis of data from dose titration trials, data from one simulation data set was used for plotting the observed dose–response relationship based on various data subsets and summaries. The same data set, and a re-estimated true model based on this data set, was used to illustrate the performance of some standard goodness-of-fit plots and to demonstrate how a meaningful visual predictive check (VPC)[23] may be performed under dose titration. The goodness-of-fit plots were plots of observed SMPG versus population and individual predictions of SMPG, plots of conditional weighted residuals (CWRES) versus population predictions and time, a QQ-plot of CWRES, and a density plot comparing the distribution of CWRES to the standard normal distribution. The VPC was performed by plotting the means and 5th and 95th percentiles of the observed doses and SMPGs by time on top of 95% confidence intervals for the means and 5th and 95th percentiles based on 1000 simulations of the trial with the estimated model. As part of the simulations, titration as defined by the titration algorithm was also simulated. R Version 3.5.3 was used for this.

RESULTS

Graphical data exploration

Data from one simulation data set for the rich-data scenario with no dropout and the linear dose–response model without disease progression is shown in Figure 2. Figure 2a,b shows the time courses of insulin dose and SMPG and illustrates how observed SMPG is gradually decreasing toward the target range of 3.9–6 mmol/L as the insulin dose is gradually increased. Figure 2c shows the observed dose–response relationship based on pooled data from all subjects at all timepoints, Figure 2d shows the observed dose–response relationship based on the medians of all doses and SMPG values over time for each subject, Figure 2e shows the observed dose–response relationship based on the end-of-trial dose and SMPG for each subject, and Figure 2f shows the true dose–response relationship used for simulation.

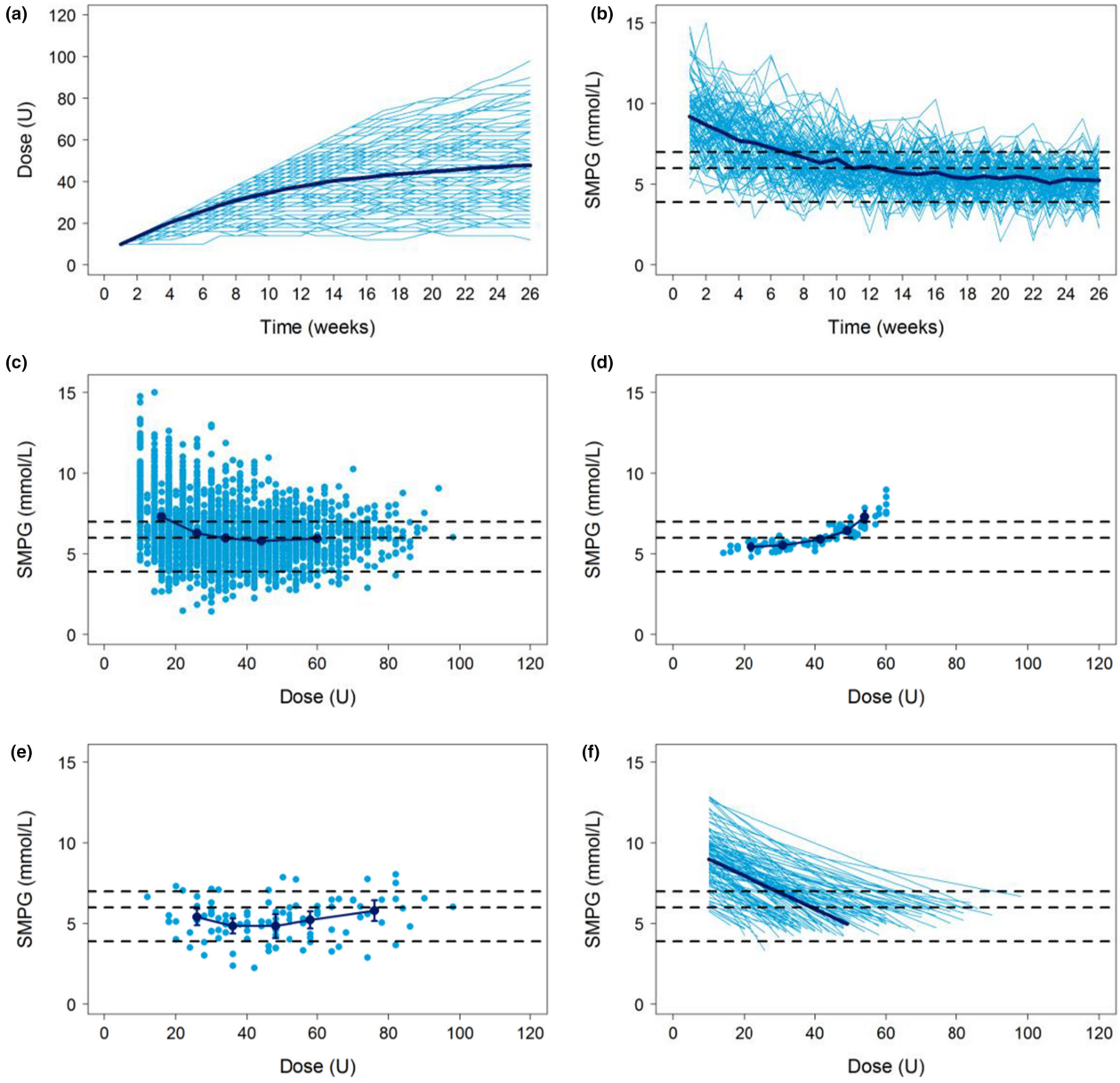

FIGURE 2 Simulated data from one example data set illustrating the difficulty of visualizing the underlying dose–response (D–R) relationship based on observed data obtained under dose titration. (a) Time course of insulin dose. (b) Time course of self-measured plasma glucose (SMPG). (c) Observed D–R based on pooled data from all individuals at all timepoints indicates the right direction of the D–R relationship at low doses but wrongly shows that it flattens out at high doses. (d) Observed D–R based on medians of all doses and SMPG values over time for each individual wrongly shows a flat D–R relationship at low doses and indicates the wrong direction of the D–R relationship at high doses. (e) Observed D–R based on end-of-trial doses and SMPG values wrongly shows a flat D–R relationship. (f) True dose–response curves used for simulation. Light blue points/lines are individual subject values/profiles, and dark blue points/lines are (population) mean values/profiles (in five equally sized bins for c–e).

Model discrimination performance

Table 2 shows the outcome of simulation and estimation in terms of discrimination between alternative models as assessed by means of empirical Type 1 and Type 2 error rates for the rich-data scenario across the five missing data scenarios. Table S1 shows similar results for the limited-data scenario.

Estimation performance

Table 3 shows the outcome of simulation and estimation in terms of relative bias, relative RMSE, and coverage probability for the estimated parameters for the rich-data scenario across the five missing data scenarios for all four variants of the simulation model. Table S2 shows similar results for the limited-data scenario.

TABLE 2 Type 1 and Type 2 error rates for model discrimination tests in the rich-data scenario with 100 subjects for 26 weeks

Missing data scenario	Model discrimination test	Type 1 (false positive) error rate (%)	Type 2 (false negative) error rate (%)
No dropout	Linear vs. nonlinear	2.1	0.0
	Linear vs. linear + progression	4.0	0.0
	Nonlinear vs. nonlinear + progression	4.2	0.0
	Linear + progression vs. nonlinear + progression	1.2	0.0
15% random dropout	Linear vs. nonlinear	2.6	0.0
	Linear vs. linear + progression	4.6	0.0
	Nonlinear vs. nonlinear + progression	4.8	0.0
	Linear + progression vs. nonlinear + progression	1.5	0.0
30% random dropout	Linear vs. nonlinear	2.2	0.0
	Linear vs. linear + progression	4.3	0.0
	Nonlinear vs. nonlinear + progression	4.4	0.0
	Linear + progression vs. nonlinear + progression	1.7	0.0
15% nonrandom dropout	Linear vs. nonlinear	0.6	0.0
	Linear vs. linear + progression	1.7	0.0
	Nonlinear vs. nonlinear + progression	3.7	0.0
	Linear + progression vs. nonlinear + progression	0.1	0.0
30% nonrandom dropout	Linear vs. nonlinear	0.3	0.2
	Linear vs. linear + progression	1.6	0.0
	Nonlinear vs. nonlinear + progression	3.7	0.0
	Linear + progression vs. nonlinear + progression	0.0	0.0

Parameter a represents the SMPG without insulin treatment, parameter b represents the glucose-lowering effect of one unit of insulin in the linear variants of the model ($c = 0$) and is equivalent to Emax/ED50 in the classic Emax model in the nonlinear variants of the model ($c > 0$), parameter c describes the degree of deviation from a linear dose–response relationship and is for the nonlinear variants of the model equivalent to 1/ED50, parameter d describes disease progression over time.

Model diagnostics

For illustrative purposes, Figure 3 shows standard goodness-of-fit plots based on the data from the simulation data set in Figure 2 and the estimated model based on this data set, and Figure 4 shows a VPC for dose and SMPG based on the same model and data set.

DISCUSSION

Our simulation study shows that development and estimation of dose–response models may also be feasible under dose titration if the data are analyzed using a mixed-effects model that accounts for the fact that dose–response relationships differ between individuals.

In the rich-data scenario with 100 subjects for 26 weeks, and across all missing data scenarios we considered, model discrimination performance was high with Type 1 error rates below the nominal 5% and Type 2 error rates very close to zero in all cases (Table 2). The low Type 2 error rate corresponds to a model discrimination power close to 100% in this scenario. In the limited-data scenario with 50 subjects for 16 weeks (Table S1), model discrimination power was 75%–95% for the missing data scenarios with no or random dropout and 60%–90% for the missing data scenarios with nonrandom dropout. Importantly, the Type 1 error rate was still below the nominal 5%. These results show that with an appropriate sample size and missing data missing at random, stepwise dose–response model development based on data obtained under dose titration is not subject to model selection bias in our example. We particularly note that we can correctly distinguish between linear and nonlinear dose–response relationships and that we are able to correctly identify if a time effect due to disease progression is present. A specific risk with

TABLE 3 Estimation performance statistics for re-estimation with the true model in the rich-data scenario with 100 subjects for 26 weeks

Missing data scenario	Model	*a*	*b*	*c*	*d*	BSV *a*	BSV *b*	BSV correlation *a*–*b*	Σ
No dropout	Linear	0.4 (2.2) [94.4]	3.0 (6.8) [92.6]	–	–	−3.0	−3.4	−9.1	−0.4
	Nonlinear	0.7 (2.9) [94.9]	4.4 (14.2) [95.6]	1.3 (20.7) [93.4]	–	−4.7	−3.8	−9.8	−0.4
	Linear + progression	0.4 (2.2) [94.9]	2.4 (7.6) [94.7]	–	2.2 (11.2) [94.8]	−1.7	−1.8	−3.0	−0.4
	Nonlinear + progression	0.4 (2.7) [94.8]	2.3 (11.4) [95.7]	−1.3 (19.4) [91.3]	3.1 (12.8) [94.7]	−2.6	−2.0	−3.6	−0.5
15% random dropout	Linear	0.3 (2.2) [94.6]	3.1 (7.0) [92.9]	–	–	−3.1	−4.0	−9.8	−0.4
	Nonlinear	0.7 (3.0) [95.1]	5.0 (15.0) [95.4]	1.9 (22.1) [93.0]	–	−5.0	−4.6	−10.9	−0.4
	Linear + progression	0.3 (2.2) [94.8]	2.7 (8.0) [94.5]	–	2.6 (12.0) [94.6]	−1.7	−2.3	−3.6	−0.4
	Nonlinear + progression	0.4 (2.7) [94.6]	2.8 (12.0) [95.3]	−1.2 (20.6) [92.1]	3.6 (13.6) [94]	−2.9	−2.6	−4.8	−0.5
30% random dropout	Linear	0.3 (2.2) [94.6]	3.4 (7.2) [93.2]	–	–	−3.1	−4.4	−10.3	−0.4
	Nonlinear	0.8 (3.0) [94.5]	5.7 (15.7) [95.6]	2.6 (23.1) [93.9]	–	−5.2	−5.3	−11.8	−0.4
	Linear + progression	0.3 (2.2) [94.6]	3.1 (8.3) [94.2]	–	3.1 (12.9) [94.6]	−1.8	−2.8	−4.3	−0.5
	Nonlinear + progression	0.5 (2.8) [95.3]	3.3 (12.5) [96.0]	−1.3 (21.5) [91.7]	4.2 (14.6) [93.2]	−3.3	−3.1	−5.8	−0.6
15% nonrandom dropout	Linear	0.9 (2.3) [92.6]	7.4 (9.8) [79.8]	–	–	−3.8	−11.4	−12.0	−0.3
	Nonlinear	0.3 (2.9) [95.4]	2.0 (13.5) [95.5]	−9.4 (22.5) [87.2]	–	−6.7	−9.4	−17.5	−0.4
	Linear + progression	0.7 (2.3) [93.3]	5.6 (9.4) [89.8]	–	−0.7 (11.1) [95.9]	−2.5	−9.3	−5.1	−0.3
	Nonlinear + progression	0.3 (2.7) [95.5]	2.1 (11.4) [95.7]	−6.6 (20.5) [87.1]	1.8 (13.1) [94.9]	−3.9	−7.8	−8.2	−0.5
30% nonrandom dropout	Linear	1.2 (2.5) [90.5]	9.7 (11.7) [70.1]	–	–	−4.6	−12.1	−14.7	−0.3
	Nonlinear	0.1 (2.9) [95.7]	0.2 (13.6) [93.4]	−16.2 (26.8) [82.3]	–	−8.1	−9.1	−22.6	−0.4
	Linear + progression	0.9 (2.4) [92.5]	7.1 (10.5) [87.8]	–	−2.4 (11.6) [95.3]	−3.1	−9.9	−7.6	−0.2
	Nonlinear + progression	0.1 (2.7) [95.3]	0.2 (11.1) [94.7]	−13.8 (23.8) [81.2]	3.4 (13.9) [93.7]	−5.4	−7.8	−13.4	−0.5

Note: Results are shown as follows: relative bias in percent (relative root mean square error in percent) [coverage probability]. Parameter *a* represents the SMPG without insulin treatment, parameter *b* represents the glucose-lowering effect of one unit of insulin in the linear variants of the model ($c = 0$) and is equivalent to Emax/ED50 in the classic Emax model in the nonlinear variants of the model ($c > 0$), parameter *c* describes the degree of deviation from a linear dose–response relationship and is for the nonlinear variants of the model equivalent to 1/ED50, parameter *d* describes disease progression over time.

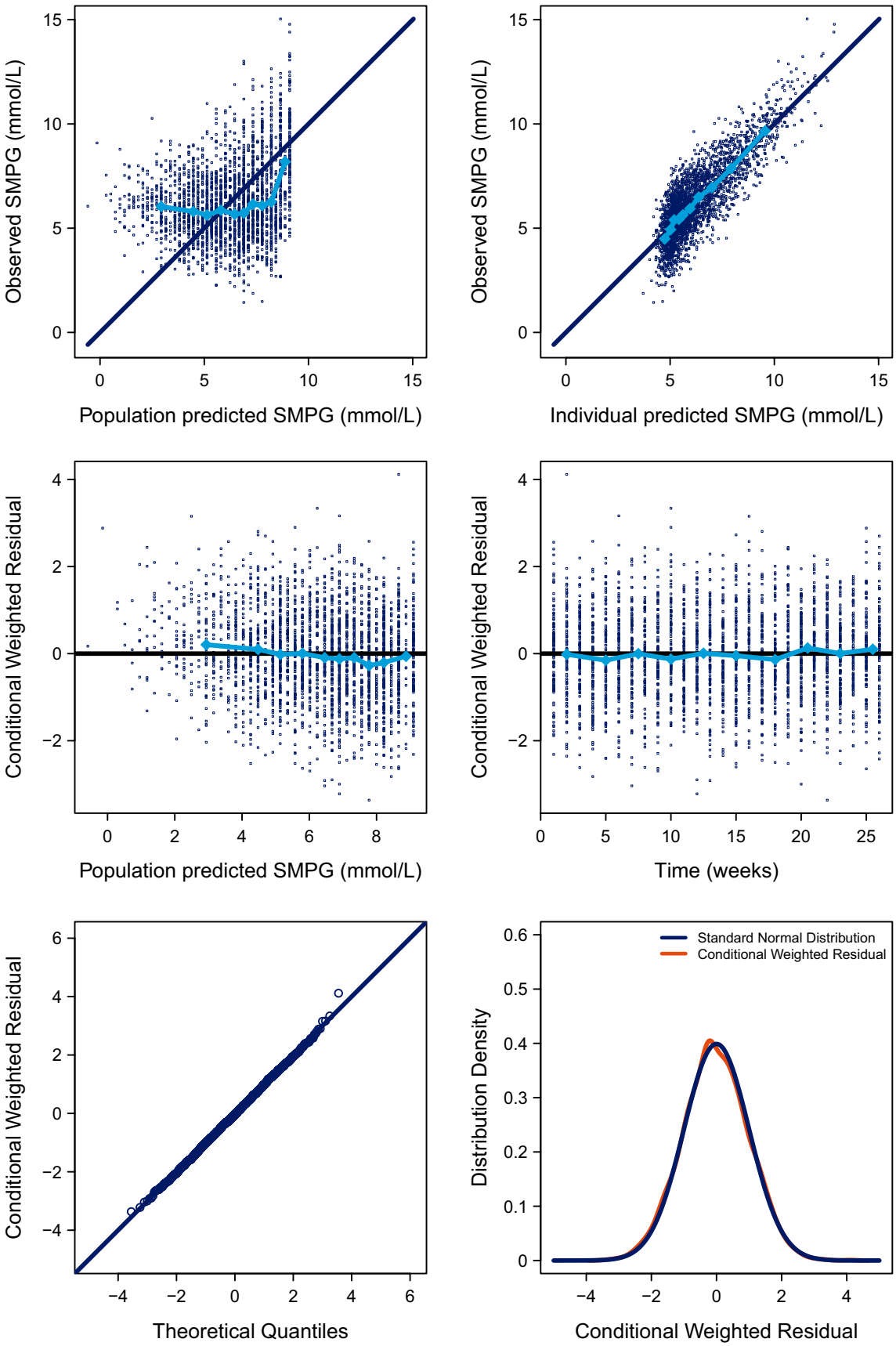

FIGURE 3 Standard goodness-of-fit plots based on the example data set in Figure 2 and the re-estimated true model based on this data set. Top row: observed self-measured plasma glucose (SMPG) values versus population and individual predictions of SMPG. Middle row: conditional weighted residuals versus population predictions and time. Bottom row: QQ-plot of conditional weighted residuals and density plot comparing the distribution of CWRES to the standard normal distribution.

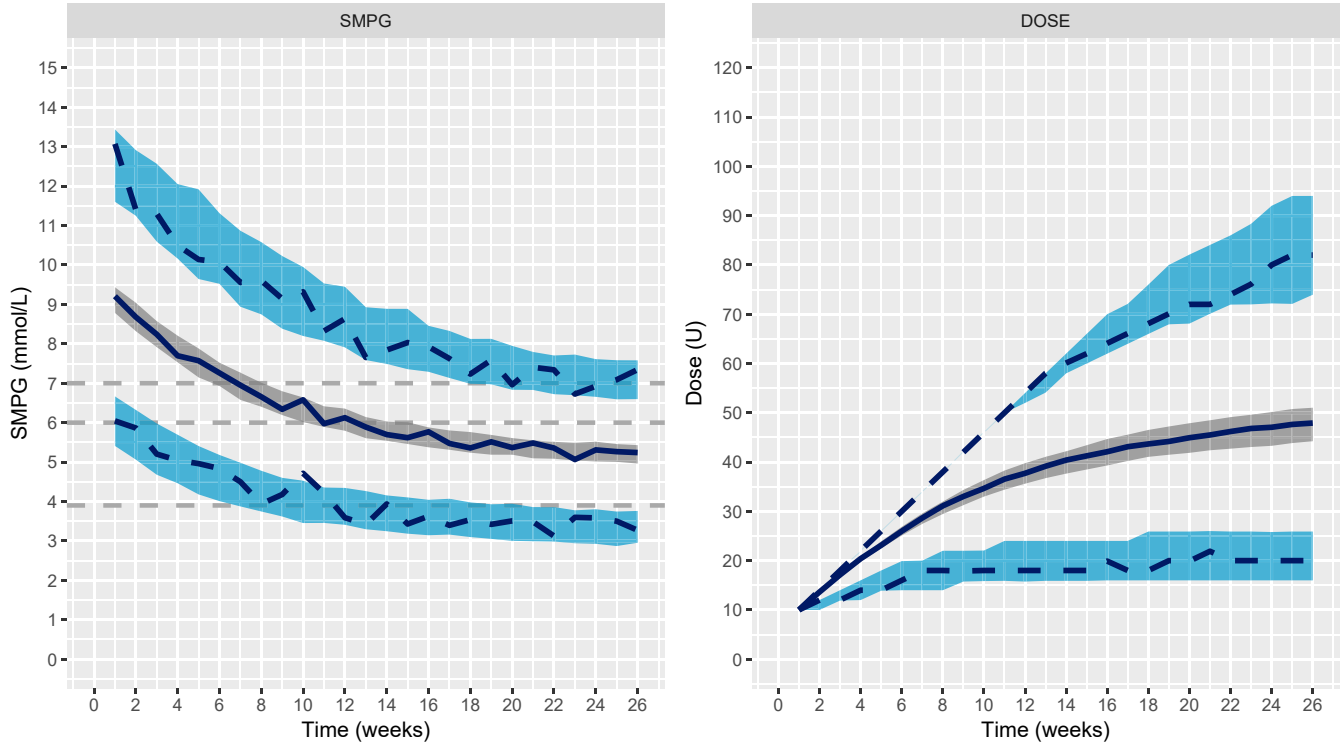

FIGURE 4 Visual predictive check based on the example data set in Figure 2 and the re-estimated true model based on this data set. Lines are means and 5th and 95th percentiles of observed data. Shaded areas are 95% confidence intervals for the same means and percentiles based on 1000 simulations. SMPG, self-measured plasma glucose.

dose–response analysis under dose titration, mentioned in regulatory guidelines, is confounding of dose and time effects,[8] but we see no evidence of that in our example and would only expect complete confounding in very rare cases.

In terms of estimation performance, the rich-data scenario had low levels of bias for all parameters in all models in the missing data scenarios with no or random dropout (Table 3). For the fixed-effect parameters and random-effect variances, relative bias was less than 6%, and for the random-effect correlation parameter, relative bias was less than 12%. Relative RMSE was less than 25%, and coverage probabilities were close to the nominal 95%. In the missing data scenarios with nonrandom dropout, relative bias was higher for a few parameters, relative RMSE was similar, and coverage probabilities were lower for a few parameters. In the limited-data scenario, similar results were obtained (Table S2), but bias was slightly higher, RMSE was higher, and coverage probabilities were slightly lower. These results suggest that with an appropriate sample size and missing data missing at random, dose–response model estimation based on data obtained under dose titration is not subject to critical bias or otherwise poor estimation performance in our example.

As expected, our results furthermore show that if missing data are not missing at random, model discrimination power decreases and bias in the estimated parameters increases, stressing the importance of accounting for any nonrandomly missing data by explicitly modeling the missingness mechanism, if possible, or using an appropriate imputation method.

Our study also highlights some pitfalls associated with applying standard approaches to graphical data exploration and model diagnostics for dose–response data obtained under dose titration. Specifically, it shows how difficult it is to visualize the underlying dose–response relationship based on observed data when it differs between individuals who as a result are titrated to different end-of-trial doses (Figure 2). A plot based on pooled data from all individuals at all timepoints (Figure 2c) may indicate the right direction of the dose–response relationship at low doses but would tend to wrongly show that it flattens out at high doses; a plot based on the averages of all doses and responses over time for each individual (Figure 2d) would also tend to show a flat dose–response relationship and may indicate the wrong direction, as observed in the recent article on the titration paradox,[6] and a plot based on end-of-trial data only (Figure 2e) would tend to show a completely flat dose–response relationship. Apart from individual-specific plots, we are not aware of simple graphical data exploration methods for appropriately visualizing the dose–response relationship using observed data. In terms of standard goodness-of-fit plots (Figure 3), our study shows that these work reasonably well under

dose titration, except for the plot of observed responses versus population-predicted responses, which may wrongly indicate that the model is mis-specified although the model structure is correct and the parameter estimates are unbiased. This is due to different realized dose ranges across individuals, where responses are underpredicted by the population mean for good responders receiving low doses, and responses are overpredicted by the population mean for poor responders receiving high doses. In terms of the VPC (Figure 4), our study shows that it is possible to produce a meaningful VPC under dose titration, if, in the VPC simulations, titration as defined by the titration algorithm is also simulated and plots are made for both dose and response.

In our study, we applied a mixed-effects model for estimation to account for the different dose–response relationships in different individuals, but we would expect the results to be similar for a two-stage approach, where the parameters are initially estimated for each individual separately and then summarized provided that sufficient data are available for each individual. Only when data are naïvely pooled across individuals, such as in Figure 2c–e, is the titration paradox observed.

Although our simulation study focused on dose–response modeling, we expect the results and conclusions to be the same for exposure–response modeling if there is a clear causal relationship between the response and the exposure level at a given timepoint and the exposure level at that timepoint is either measured or can be predicted based on the current and previous dose levels using a pharmacokinetic model. If the response is delayed or develops slowly over time such that there is not a clear causal relationship between the response and the exposure level at a given timepoint, more advanced methods will be needed. If a deterministic dose-modification algorithm has been used, such as in our simulation example and commonly in well-controlled, randomized clinical trials, DLME[12] can be applied, or a longitudinal mixed-effects dose-exposure-response model may be applied, where the time course of the response is explicitly modeled. If the dose-modification algorithm is not well defined, such as in observational studies, MSM with IPTW[10,11] can be considered because this method implicitly includes the development of a model for the dose-modification process. However, obtaining unbiased results with the MSM approach may be challenging.[11]

The results of our study show that dose/exposure–response modeling should not be disregarded upfront out of concerns for selection bias if dose titration is applied in a clinical trial or drug development program. Although it cannot be concluded in general that using a population approach will always ensure proper model discrimination and unbiased parameter estimates, we have shown that if there is a clear causal relationship between the response and the dose/exposure level at a given timepoint, it will often be possible to use such an approach to develop, estimate, and appropriately qualify a dose/exposure–response model, even though the data were obtained under dose titration. The trials amenable to using such an approach, provided the assumption of causality between dose/exposure and response is fulfilled and a well-defined, deterministic titration algorithm is applied, would typically be exploratory or confirmatory trials with sufficient duration relative to the frequency of dose titration to ensure that the majority of subjects reach the desired response and therefore have observations of response across a range of different doses. In practice, a prospective evaluation, in line with the analysis presented here, of the feasibility of using such an approach in a specific case would be recommended, as there will also be cases where the specific trial design or titration algorithm applied does not produce data that adequately support the use of a population approach. For example, a high prevalence of nonresponders or individuals who do not titrate due to adequate response at the initial dose could give imbalances in the data that would lead to bias. Likewise, a high prevalence of extreme responders and other deviations from a unimodal, normal-type distribution of responses could lead to bias unless the specific distribution can be explicitly modeled.

The possible applications of a dose/exposure–response model based on data obtained under dose titration are many. If developed from exploratory trial data, the model could be applied to perform simulations to develop, evaluate, and optimize the titration algorithm to be used in confirmatory trials, for example, in terms of titration target, starting dose, and dose increments, thus reducing the risks associated with studying a novel titration algorithm in pivotal trials and potentially improving on the current underuse of titration in such trials and the resulting lack of titration guidance in drug labeling.[2,5] If applied to a larger pool of confirmatory trial data, the model could be applied for covariate analysis to investigate the impact of demographics and baseline characteristics to facilitate further individualization of titration algorithms and optimization of individual benefit–risk balance.

Based on our study, we conclude that the use of dose titration does not by definition cause a "titration paradox" and issues with selection bias when longitudinal data from such trials are analyzed. If, at every timepoint in the course of a trial, there is a clear causal relationship between the dose/exposure level and the response, and if a population approach is used, it will in many cases be possible to make appropriate inferences from the data obtained under dose titration.

AUTHOR CONTRIBUTIONS

N.R.K. designed and performed the research, analyzed the data, and wrote the manuscript. H.A. designed the research and wrote the manuscript.

ACKNOWLEDGMENT

The authors would like to acknowledge Søren Andersen, Biostatistics, Novo Nordisk A/S, for providing statistical input to the research and for reviewing the manuscript.

FUNDING INFORMATION

No funding was received for this work.

CONFLICT OF INTEREST

Niels Rode Kristensen and Henrik Agersø are employees and shareholders of Novo Nordisk A/S.

REFERENCES

1. Peck RW. Precision dosing: an industry perspective. *Clin Pharmacol Ther*. 2021;109:47-50.
2. Neely M, Onufrak N, Scheetz MH, et al. Supporting precision dosing in drug labeling. *Clin Pharmacol Ther*. 2021;109:37-41.
3. Vinks AA, Peck RW, Neely M, Mould DR. Development and implementation of electronic health record–integrated model-informed clinical decision support tools for the precision dosing of drugs. *Clin Pharmacol Ther*. 2020;107:129-135.
4. Chen C. The remarkable therapeutic potential of response-based dose individualisation in drug trials and patient care. *Drug Discov Today*. 2018;23:1463-1468.
5. Schuck RN, Pacanowski M, Kim S, Madabushi R, Zineh I. Use of titration as a therapeutic individualization strategy: an analysis of Food and Drug Administration-approved drugs. *Clin Transl Sci*. 2019;12:236-239.
6. Schnider TW, Minto CF, Filipovic M. The drug titration paradox: correlation of more drug with less effect in clinical data. *Clin Pharmacol Ther*. 2021;110:401-408.
7. *ICH HARMONISED TRIPARTITE GUIDELINE* – E4: Dose-Response Information to Support Drug Registration. 1994.
8. US Food and Drug Administration. Guidance for industry: exposure-response relationships - study design, data analysis, and regulatory applications. 2017.
9. Overgaard RV, Ingwersen SH, Tornøe CW. Establishing good practices for exposure-response analysis of clinical endpoints in drug development. *CPT Pharmacometrics Syst Pharmacol*. 2015;4:565-575.
10. Lipkovich I, Adams DH, Mallinckrodt C, Faries D, Baron D, Houston JP. Evaluating dose response from flexible dose clinical trials. *BMC Psychiatry*. 2008;8:3.
11. Lipkovich I, Mallinckrodt CH, Faries DE. The challenges of evaluating dose response in flexible-dose trials using marginal structural models. *Pharmaceut Statist*. 2012;11:485-493.
12. Xu X, Yuan M, Nandy P. Analysis of dose-response in flexible dose titration clinical studies. *Pharmaceut Stat*. 2012;11:280-286.
13. Funatogawa I, Funatogawa T. Dose-response relationship from longitudinal data with response-dependent dose modification using likelihood methods. *Biom J*. 2012;54:494-506.
14. Riddle M, Rosenstock J, Gerich J. The treat-to-target trial: randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients. *Diabetes Care*. 2003;11:3080-3086.
15. Holt RIG, DeVries JH, Hess-Fischl A, et al. The management of type 1 diabetes in adults. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of diabetes (EASD). *Diabetologia*. 2021;64:2609-2652.
16. Davies MJ, D'Alessio DA, Fradkin J, et al. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of diabetes (EASD). *Diabetes Care*. 2018;2018(41):2669-2701.
17. Lantus®. https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/021081s072lbl.pdf
18. Tresiba®. https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/203314lbl.pdf
19. Holford N. Clinical pharmacology = disease progression + drug action. *Br J Clin Pharmacol*. 2013;79:18-27.
20. Ollerton RL, Playle R, Ahmed K, Dunstan FD, Luzio SD, Owens DR. Day-to-day variability of fasting plasma glucose in newly diagnosed type 2 diabetic subjects. *Diabetes Care*. 1999;22:394-398.
21. Heise T, Nosek L, Bøttcher SG, Hastrup H, Haahr H. Ultra-long-acting insulin degludec has a flat and stable glucose-lowering effect in type 2 diabetes. *Diabetes Obes Metab*. 2012;14:944-950.
22. Heise T, Hövelmann U, Nosek L, Hermanski L, Bøttcher S, Haahr H. Comparison of the pharmacokinetic and pharmacodynamic profiles of insulin degludec and insulin glargine. *Expert Opin Drug Metab Toxicol*. 2015;11(8):1193-1201.
23. Holford N. The Visual Predictive Check – Superiority to Standard Diagnostic (Rorschach) Plots. Page 14, Abstr 738. https://www.page-meeting.org/?abstract=738. 2005.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kristensen NR, Agersø H. Dose/exposure–response modeling in dose titration trials: Overcoming the titration paradox. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:1592-1603. doi:10.1002/psp4.12863