ORIGINAL ARTICLE

Population Exposure-Response Modeling of Naloxegol in Patients With Noncancer-Related Pain and Opioid-Induced Constipation

N Al-Huniti[1], JC Nielsen[2], MM Hutmacher[2], J Lappalainen[1], K Cantagallo[1] and M Sostek[1]

Naloxegol is a polyethylene glycol derivative of naloxone approved in the US as a once-daily oral treatment for opioid-induced constipation (OIC) in adults with chronic noncancer pain. Population exposure–response models were constructed based on data from two phase III studies comprising 1,331 adults with noncancer pain and OIC. In order to characterize the protocol-defined naloxegol responder rate, the number of daily spontaneous bowel movements (SBMs) was characterized by a longitudinal ordinal nonlinear mixed-effects logistic regression dose–response model, and the incidence of diary entry discontinuation was described by a time-to-event model. The mean number of SBMs per week increased with increasing naloxegol dose. The predicted placebo-adjusted responder rates (90% confidence interval) were 10.4% (4.6–13.4%) and 11.1% (4.8–14.4%) for naloxegol 12.5 and 25 mg/day, respectively. Model-predicted response to naloxegol was influenced by the baseline SBM frequency and characteristics of the opioid treatment.
CPT Pharmacometrics Syst. Pharmacol. (2016) **5**, 359–366; doi:10.1002/psp4.12099; published online 20 July 2016.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
☑ In phase II and III clinical studies, naloxegol 25 mg/day significantly increased the responder rate compared with placebo in patients with opioid-induced constipation taking opioids for noncancer pain.
• WHAT QUESTION DID THIS STUDY ADDRESS?
☑ In a population exposure–response analysis, the number of spontaneous bowel movements (SBMs) per day and the diary entry discontinuation (DED) rate were modeled as predictors for responder rate. The influence of prespecified covariates was investigated.
• WHAT THIS STUDY ADDS TO OUR KNOWLEDGE
☑ The number of SBMs was characterized by a non-linear mixed-effects logistic regression dose–response model. DED rate was described by a time-to-event model. Predicted responder rates were similar for naloxegol 12.5 and 25 mg/day. Response was predicted to be higher in patients with lower baseline SBM frequency (i.e., more severe constipation) and those taking a strong opioid.
• HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS
☑ Naloxegol appeared to provide larger treatment effect in patients with low baseline SBM, suggesting that the 25 mg dose may provide additional benefit in patients with more severe disease.

Opioid-induced constipation (OIC) is a common side effect associated with opioid treatment. The estimated prevalence of OIC is 15–90% in patients receiving opioids for non-cancer pain,[1] most of whom regard OIC as the most bother-some side effect, with a negative impact on quality of life.[2]

Naloxegol is a polyethylene glycol derivative of naloxone that has been approved in the US as an oral, once-daily treatment for OIC in adults with chronic noncancer pain.[3] Since naloxegol has limited ability to cross the blood–brain barrier, the central analgesic properties of opioid agonists are maintained.[4] Laxatives do not address the underlying causes of OIC and up to 46% of patients do not achieve the desired treatment outcome.[5] Naloxegol, by binding to μ-opioid receptors within the gastrointestinal tract, targets the underlying causes of OIC, i.e., reduced gastrointestinal motility, hypertonicity, and increased fluid absorption result-ing from long-term opioid treatment.

Randomized, double-blind, placebo-controlled phase II and III studies have demonstrated the efficacy of naloxegol in patients with OIC and noncancer pain. Patients used an eDiary to record spontaneous bowel movements (SBMs), defined as bowel movements that occurred in the previous 24 hours without the use of rescue laxatives. In a phase II study ($N = 207$) naloxegol 25 mg/day for 4 weeks was sig-nificantly more effective than placebo in increasing the num-ber of SBMs per week over baseline ($P = 0.0022$).[4] Two identical, multicenter, double-blind, phase III studies (studies 4 (KODIAC 4) and 5 (KODIAC 5); ClinicalTrials.gov identi-fiers NCT01309841 and NCT01323790, respectively) ran-domized >1,350 patients to naloxegol 12.5 or 25 mg/day or placebo.[6] The primary efficacy endpoint was responder rate over the 12-week treatment period, defined as ≥3 SBMs/week with ≥1 SBM/week increase over baseline for ≥9 of 12 weeks, including ≥3 SBMs/week during the last 4 weeks

[1]AstraZeneca Pharmaceuticals, Waltham, Massachusetts, USA; [2]Ann Arbor Pharmacometrics Group, Ann Arbor, Michigan, USA. *Correspondence to: N Al-Huniti (Nidal.Al-Huniti@astrazeneca.com)

of treatment. In both studies, naloxegol 25 mg/day significantly increased the responder rate compared with placebo (naloxegol, 40–44%; placebo, 29%).[6]

The present study concerns a population exposure–response analysis conducted using data from studies 4 and 5. A longitudinal ordinal logistic regression model was developed to characterize the relationship between naloxegol dose and the daily number of SBMs. In addition, a model for the time to diary entry discontinuation (DED) was developed, and the two models were used together to predict responder rate.

METHODS
Subjects and study design
Studies 4 and 5 had an identical design: phase III, multicenter, double-blind, randomized, placebo-controlled, parallel group studies of 12.5 and 25 mg/day naloxegol. The studies comprised a 2-week screening period; 2-week OIC diagnosis confirmation (baseline) period; 12-week treatment period; and a follow-up visit 2 weeks after the last dose of study drug.

Eligible patients were adults receiving stable maintenance opioids (total daily dose 30–1,000 mg of oral morphine or equivalent) for ≥ 4 weeks for noncancer-related pain, with self-reported <3 SBMs per week and active OIC symptoms. Key exclusion criteria included opioid treatment for cancer-related pain; history of cancer within past 5 years (apart from basal cell and squamous cell skin cancer); and other conditions or treatment related to the gastrointestinal tract. See **Supplementary Information** for further details.

The studies were approved by the local Ethics Committees and were performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients.

Model development
Model development was based on the intention-to-treat population, defined as patients who were randomized and received at least one dose of study drug. The number of SBMs per day was incorporated as a time-varying covariate for the DED analysis. Missing SBM values were imputed using the last-observation-carried-forward approach. For both analyses, missing values were not imputed for the other covariates; patients with missing data were excluded.

Nonlinear mixed-effects modeling methodology was implemented using the Laplacian estimation method in NONMEM (v. 7.2).[7] SAS (Cary, NC, v. 9.3) was used for postprocessing of NONMEM output and for simulation purposes. S-plus (v. 8.2, TIBCO Software, Palo Alto, CA, 2010) was used for graph construction.

The SBM and DED models were developed using a similar approach. Initially, baseline and nondrug components were assessed using baseline (SBM model only) and placebo data. This strategy facilitated development of baseline and nondrug effect components while avoiding the potential impact of drug model misspecification. Following identification of a suitable baseline and nondrug effect model, the effect of naloxegol exposure was evaluated by exploring a drug effect model. Steady-state concentrations were used due to the lack of any significant naloxegol accumulation following repeated dosing. Individual predicted naloxegol average concentration (C_{avg}), maximum concentration (C_{max}), and area under the concentration time curve for the 24 hours dosing interval from the population pharmacokinetic (PK) model[8] (**Supplementary Information**) were evaluated. Prediction of these varied daily because of the time-varying covariates identified in the population PK analysis (e.g., concurrent use of a CYP3A4 inhibitor).

After the base structural model was identified, a full model was constructed in which prespecified covariates (**Supplementary Tables S1, S2**) were added to the model simultaneously. Covariates were incorporated in a linear fashion on model parameters that were allowed to be positive or negative. Power models (log-linear of the log transformed covariate value) were used for those parameters constrained to be >0. Opioid potency was considered to be interchangeable with opioid dose. As there is no definitive standard for interconversion of opioid doses, the doses were converted to oral morphine equivalents.

Model development was guided by successful convergence and calculation of standard errors; reductions in objective function values (OFVs) for hierarchical models; and overall goodness-of-fit. The stability of the models was evaluated throughout development.

Once a stable full model was completed, a selection procedure was implemented in order to identify a parsimonious covariate model. The Wald Approximation Method (WAM) procedure[9] was used to identify a subset of reduced models relative to the full model. The top 15 models ranked according to Maximizing Schwarz's Bayesian Criterion (SBC) were fitted using NONMEM to calculate the actual SBC. The final parsimonious model was selected based on the maximum NONMEM-based SBC. High rank order correlation between the Wald- and NONMEM-based SBC statistics was used to verify good performance of the WAM algorithm. Finally, the predictive performance of the reduced covariate model was evaluated prior to selection of the final model.

Daily spontaneous bowel movement model
With regard to the nondrug effects model, evaluation of each unique transition was not undertaken, since a large number of parameters would be required. Since no consistent trends with time were noted during the baseline period, changes in the probability of a particular SBM count over time were initiated at the start of randomized treatment.

For development of the base structural model (**Supplementary Table S3**), a nonlinear mixed-effects logistic regression model was fitted to the daily SBM data to describe the cumulative probability of having a daily SBM count $\geq m$ ($m = 1, 2, 3, 4, 5, 6$). SBM count data were truncated at six since there were very few SBM counts greater than six. This approach provided greater flexibility by modeling the distribution in counts as compared with using count models. The general form of the SBM model is:

$$logit\left\{P\left(SBM_{ij} \geq m | \eta\right)\right\} = \sum_{k=1}^{m} \beta_k + f_{nd}\left(t_{ij}\right) + f_d\left(EX_{ij}\right) + \eta_i$$

where SBM_{ij} denotes the number of SBMs for the ith patient on day j; $logit\{P(SBM_{ij} \geq m | \eta)\}$ represents the probability of

having an SBM count $\geq m$ for a given vector of subject-specific random effects with variance-covariance matrix, Ω; the β_k (k = 1, ..., 6) represent the baseline logit for the SBM count categories (note that there is no baseline logit-probability for 0 since by definition the probability that $m \geq 0$ is 1); $f_{nd}(t_{ij})$ is a function describing the nondrug effects; and $f_d(EX_{ij})$ is a function describing the drug effects related to naloxegol exposure (e.g., dose, C_{avg}).

Initial models assessing the impact of naloxegol exposure on the probability of SBM counts used either naloxegol dose or C_{avg} as the exposure measure. Prior to evaluation of covariate effects, OFVs together with other modeling criteria were used to select which exposure metric should be used for the remainder of model development.

A posterior predictive check[10] was performed on the final model. Uncertainty in the model parameter estimates was incorporated using a smoothed parametric bootstrap procedure.

Diary entry discontinuation model
The DED model was constructed using a time-to-event framework. The influence of the daily number of SBMs on the log-hazard was examined during structural model development (**Supplementary Table S5**), since efficacy and exposure (i.e., as a surrogate for safety) were anticipated to be key predictors of discontinuation. The general form of the log-hazard contained a baseline parameter (θ_o), and additive nondrug (f_{nd}) and drug (f_d) components:

$$log(\lambda(\theta; m)) = \theta_o + f_{nd}(\theta; m) + f_d(\theta; EX, m)$$

where m represents time and EX represents naloxegol exposure (i.e., either dose or average concentration).

Predictive performance and population simulations for responder rate
Predictive performance of the final DED model was evaluated by a visual predictive check. Kaplan–Meier (KM) estimates of survival curves for the observed data were compared with KM estimates of 200 simulated datasets. In order to be more conservative with regard to the adequacy of the final model, uncertainty in the parameter estimates was not included for the visual predictive check. A visual predictive check targeting the hazard was also conducted.[11]

Predictive performance of the SBM and DED models on responder rate was assessed using posterior predictive checks. The SBM model was used to simulate daily SBM counts in 500 datasets, each of which was conditioned on the complete design (i.e., day −13 to day 84) and covariates of the observed dataset. Uncertainty in the final SBM model parameter estimates was incorporated using a smoothed parametric bootstrap. The DED model was then applied to each simulated dataset to replicate the DED rate from the observed data. Uncertainty in the DED model was considered for the simulation of the incidence of DED. Responder rate was calculated for each simulated dataset by study and treatment group. The 90% prediction interval for the simulated responder rates was compared with the observed responder rate.

For the population simulations for mean responder rate, 500 sets of SBM and DED model parameters were sampled

using a smoothed parametric bootstrap procedure. Each set of parameter estimates was used to simulate a dataset with 2,000 patients in each treatment arm in order to characterize the uncertainty in the various metrics of interest. Baseline opioid potency was sampled from the empiric distribution of patients in studies 4 and 5. The population mean responder rate, placebo-corrected responder rate, and DED rate along with corresponding 90% confidence intervals (CIs) were generated for each treatment group.

RESULTS

The analyses were based on data from 1,331 adults. Demographic and baseline clinical characteristics are summarized in **Table 1**.

The SBM analysis included data recorded during the 2-week OIC confirmation (baseline) period and the 12-week treatment period (i.e., day −13 to day 84). The analysis included 108,051 daily diary entries, most of which recorded zero (55.1%) or one (40.1%) SBM.

The DED analysis only included data recorded during the 12-week treatment period (i.e., day 0 to day 84). DED was defined as the day after the last nonmissing diary entry, assuming time = 0 to be immediately prior to the first dose of study drug. Overall, 543 patients (40.8%) contributed their last diary entry prior to day 84. However, this is likely to be an overestimation of the percentage of patients who did not complete the study, since the window for the final study visit was within 3 days before or after day 84.

Daily spontaneous bowel movement model
Negative binomial and Poisson count models failed to adequately describe the SBM distribution. The best fitting Poisson model overpredicted the probability of zero SBMs and underpredicted the probability of one SBM (**Supplementary Figure S1**). Consequently, a mixed-effects ordered categorical logistic regression dose–response model was used.

Nondrug effects were assessed first using the baseline and placebo data only. Probabilities were estimated for the baseline and postbaseline periods, since there was a distinct demarcation in the mean number of SBMs after treatment. In addition, because the number of SBMs on a given day may be dependent on the number of SBMs on the previous day, Markov effects were evaluated based on whether a patient had an SBM on the previous day. Inclusion of nine Markov parameters resulted in a 213 unit decrease in the OFV; therefore, the Markov component was retained in the model. Linear, log-linear, and exponential plateau models were considered for the nondrug effect. The exponential plateau model best described placebo component to SBM count probability over time.

The effects of naloxegol exposure were assessed using linear, maximum drug effect (E_{max}), sigmoid E_{max}, and power models, and an E_{max} model was selected as the base dose–response model. Individual predicted concentration-based metrics did not improve the fit relative to the dose–response model.

The full model, which included all covariate effects simultaneously (i.e., 39 covariates), was overparameterized and

Table 1 Demographic and baseline clinical characteristics (intention-to-treat population)

Characteristic	Study 4 (KODIAC 4)			Study 5 (KODIAC 5)		
	Placebo (N = 214)	Naloxegol 12.5 mg (N = 213)	Naloxegol 25 mg (N = 214)	Placebo (N = 232)	Naloxegol 12.5 mg (N = 232)	Naloxegol 25 mg (N = 232)
Age (yr), mean (SD)	52.9 (10.0)	51.9 (10.4)	52.2 (10.3)	52.3 (11.6)	52.0 (11.0)	51.9 (12.1)
Female, *n* (%)	140 (65.4)	135 (63.4)	118 (55.1)	145 (62.5)	149 (64.2)	147 (63.4)
Race,[a] *n* (%)						
White	160 (74.8)	164 (77.0)	173 (80.8)	183 (78.9)	187 (80.6)	189 (81.5)
Black	44 (20.6)	42 (19.7)	38 (17.8)	44 (19.0)	41 (17.7)	40 (17.2)
Asian	4 (1.9)	5 (2.3)	1 (0.5)	0	1 (0.4)	0
Other	6 (2.8)	2 (0.9)	2 (0.9)	5 (2.2)	3 (1.3)	3 (1.3)
Duration of current opioid use (months), mean (SD)	39.5 (39.4)	44.4 (47.3)	44.5 (47.8)	43.0 (51.4)	48.5 (48.7)	40.9 (41.6)
Opioid dose (mg/day),[a] mean (SD)	135.6 (145.8)	139.7 (167.4)	143.2 (150.1)	119.9 (103.8)	151.7 (153.0)	136.4 (134.3)
OIC characteristics, mean (SD)						
No of SBM/week	1.4 (0.89)	1.4 (0.85)	1.3 (1.11)	1.5 (0.95)	1.6 (1.05)	1.3 (0.85)
Severity of straining score[a]	3.3 (0.78)	3.1 (0.79)	3.2 (0.84)	3.3 (0.81)	3.1 (0.82)	3.2 (0.82)
Stool consistency score[a]	2.8 (1.22)	2.9 (1.20)	2.9 (1.16)	3.0 (1.29)	3.0 (1.29)	2.8 (1.26)
Laxative use, *n* (%)						
Within previous 6 mo	177 (82.7)	184 (86.4)	181 (84.6)	197 (84.9)	189 (81.5)	194 (83.6)
Within previous 2 wk	151 (70.6)	140 (65.7)	166 (77.6)	173 (74.6)	156 (67.2)	166 (71.6)
Inadequate response to laxatives,[a] *n* (%)	118 (55.1)	115 (54.0)	117 (54.7)	121 (52.2)	125 (53.9)	124 (53.4)

Numbers of patients differed between the intention-to-treat population (641 in study K4 and 696 in study K5) and the population of patients randomly assigned to a study group (652 in study K4 and 700 in study K5) because 11 patients in study K4 and 4 patients in study K5 were found to be participating at more than one center within the program and were excluded from the intention-to-treat population. No notable between-group differences in demographic or clinical characteristics were observed; a formal statistical comparison was not performed.
From: Chey WD, Webster L, Sostek M, Lappalainen J, Barker PN, Tack J. Naloxegol for opioid-induced constipation in patients with noncancer pain; volume 370; pages 2387–2396. Copyright © (2014) Massachusetts Medical Society. Reprinted with permission from Massachusetts Medical Society.
OIC, opioid-induced constipation; SBM, spontaneous bowel movements; SD, standard deviation.
[a]Race was self-reported.
[b]This characteristic was assessed among patients in the safety-analysis set, which included all patients in the intention-to-treat population who received at least one dose of drug.
[c]Severity of straining was measured on the following scale: 1, not at all; 2, a little bit; 3, a moderate amount; 4, a great deal; and 5, an extreme amount.
[d]Stool consistency was assessed on the Bristol Stool Scale (types 1 through 7, with 1 denoting small, hard, lumpy stool, and 7 denoting watery stool).
[e]Patients with an inadequate response to laxatives were those who took laxatives in one or more laxative classes for a minimum of 4 days within 2 weeks before screening and had ratings of moderate, severe, or very severe on one or more of the four stool-symptom domains in the baseline laxative-response questionnaire.

resulted in a 106-unit reduction in the OFV. As a result, three separate models were fitted in which the covariate effects were estimated on only baseline, nondrug magnitude, or E_{max}. The magnitude of each covariate effect as well as the precision of the estimate was used to determine which covariates should be included in a working full model. The resulting working full model included the effects of: unknown baseline laxative response, age, baseline opioid daily dose and duration of opioid use on nondrug magnitude; and age, body mass index, black or Asian race, and baseline use of weak opioids on E_{max}.

The WAM did not perform well in identifying a parsimonious covariate model, judged by the poor correlation between Wald-based and NONMEM-based model rankings. Instead, a stepwise backward elimination procedure was implemented with a cutoff (in change in OFV) of 11.58 units, the penalty based on the SBC used for the WAM. All covariate effects included in the working full model were removed with the exception of baseline use of weak opioids on E_{max}.

As a result, the base structural model plus the effect of baseline use of weak opioids on E_{max} was declared the final model. The model equations and parameter estimates for the final model are provided in **Supplementary Table S4**.

The placebo (nondrug) effect was relatively small (0.189 logit units) and was achieved in ~93 days (five half-lives). The mean number of SBMs per week increased with increasing naloxegol dose. The estimated E_{max} was 1.46 logit units and the naloxegol dose resulting in 50% of the maximum effect (ED_{50}) was 18.9 mg. The ED_{50} was poorly estimated, with a 95% CI ranging from 5.70 to 62.8 mg. Visual predictive checks showed that the observed mean number of SBMs per day was consistent with the simulated data (**Figure 1**). Baseline opioid potency (weak vs. strong) was the only covariate that had a significant effect on the maximum response to naloxegol (**Figure 2**). The typical E_{max} decreased by 67% in patients who did not receive a high opioid dose relative to those who received a high opioid dose.
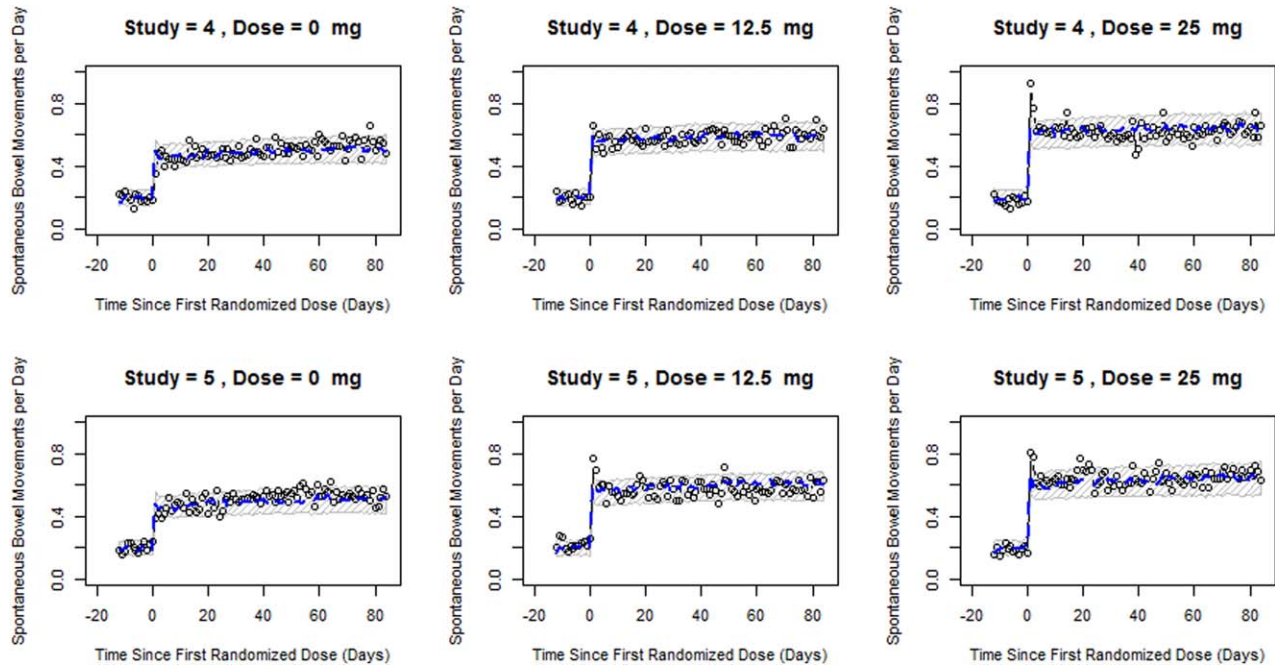
Figure 1 Visual predictive check for mean daily spontaneous bowel movements for study 4 (KODIAC 4) and study 5 (KODIAC 5).

A covariance matrix of subject-specific random effects (Ω) was estimated for the base, full, and final models. Correlations between subject-specific random effects on the postbaseline period and subject-specific random effects on the baseline period, maximum nondrug magnitude, and E_{max} were the most important. **Figure 3** illustrates that patients with lower baseline SBM values would be expected to show a smaller overall placebo effect and greater drug effect than those with higher baseline SBM values.
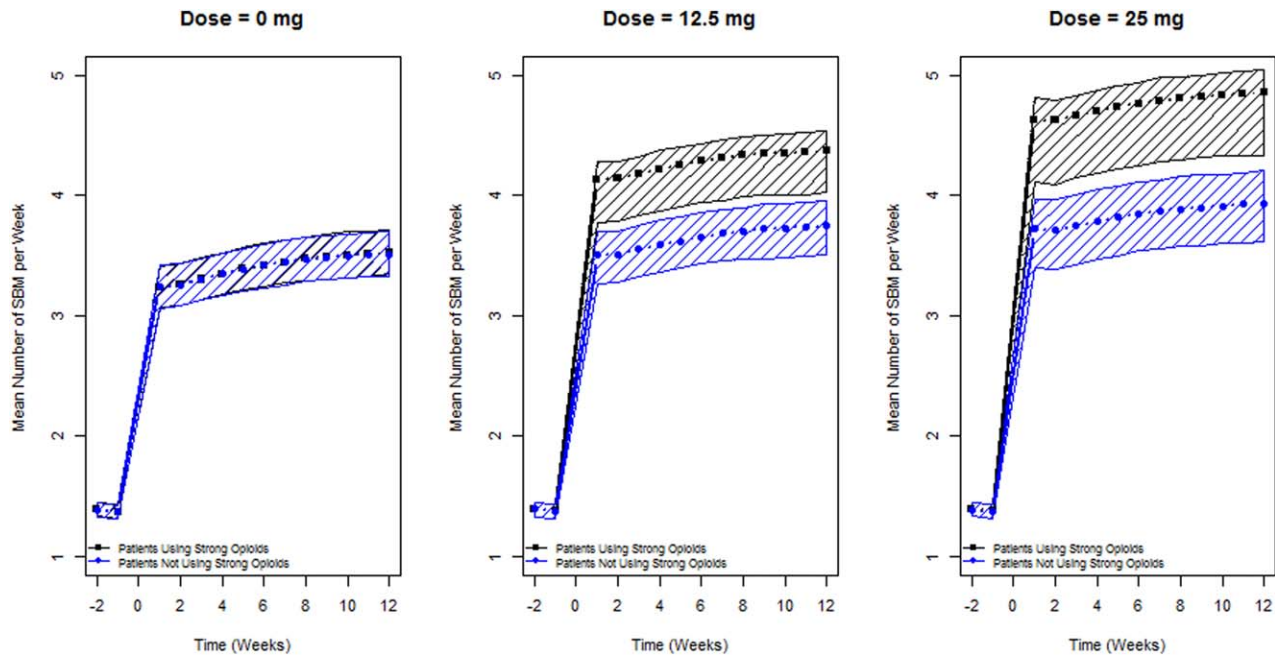
Figure 2 Illustration of impact of baseline opioid potency on maximum drug effect. Squares (shaded region) show predicted population mean (90% CI) number of SBMs per week. Black represents patients who received treatment at baseline with strong opioids; blue represents patients who did not receive treatment at baseline with strong opioids.
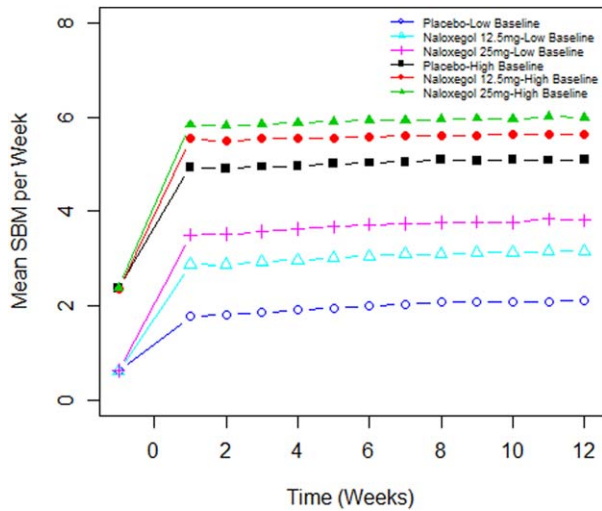
Figure 3 Illustration of correlation between subject-specific random effects. Random effects were sampled conditionally upon a low (5th percentile) and high (95th percentile) value from the estimated σ distribution for the subject-specific random effect on the baseline period.

Diary entry discontinuation model

Baseline and nondrug effects for the DED model were evaluated first in the placebo data alone (see **Supplementary Information** for equations). An observed large increase in DEDs near the end of the double-blind treatment period was attributed to the window for the final study visit (± 3 days of day 84). The increase in DEDs after day 80 reflects those patients who completed their final visit prior to day 84. An additional model component was added to address this feature of the data. The most parsimonious model tested in the placebo data was described by the following equation:

$$log(\lambda(m)) = \theta_{base} + \theta_{slp} \cdot m + [\theta_{end} \cdot (m-80)] \cdot I_{T80}$$

where θ_{base} is the parameter for the baseline log-hazard, θ_{slp} is the parameter describing the change in the log-hazard with time over the entire double-blind treatment period, θ_{end} is an additional slope parameter governing the change in the log-hazard with time near the end of the double-blind treatment period, and I_{T80} is an indicator variable set to one at times greater than or equal to 80 and zero for times less than 80. Different cutoff times starting from day 84 were evaluated, but day 80 resulted in the lowest OFV.

Following selection of a suitable nondrug model, the effects of naloxegol exposure (i.e., dose) on the log-hazard were assessed in the complete dataset. Linear and E_{max} drug models were tested initially. The E_{max} model performed better than the linear drug model (OFV decreased by 25.5 units), but the parameter estimate for ED_{50} was very small, suggesting that a simple shift for doses greater than zero would be a more parsimonious model. A model incorporating a simple step function for naloxegol doses greater than zero had a very similar OFV with one fewer parameter.

The influence of SBMs on the log-hazard was examined during structural model development, since efficacy was hypothesized to be a key predictor of study discontinuation (i.e., DED). The number of SBMs was included as a linear function on the baseline log-hazard and the slope describing the change in the log-hazard with time over the double-blind treatment period in two separate models. The OFV was reduced by less than two units in each model; therefore, the
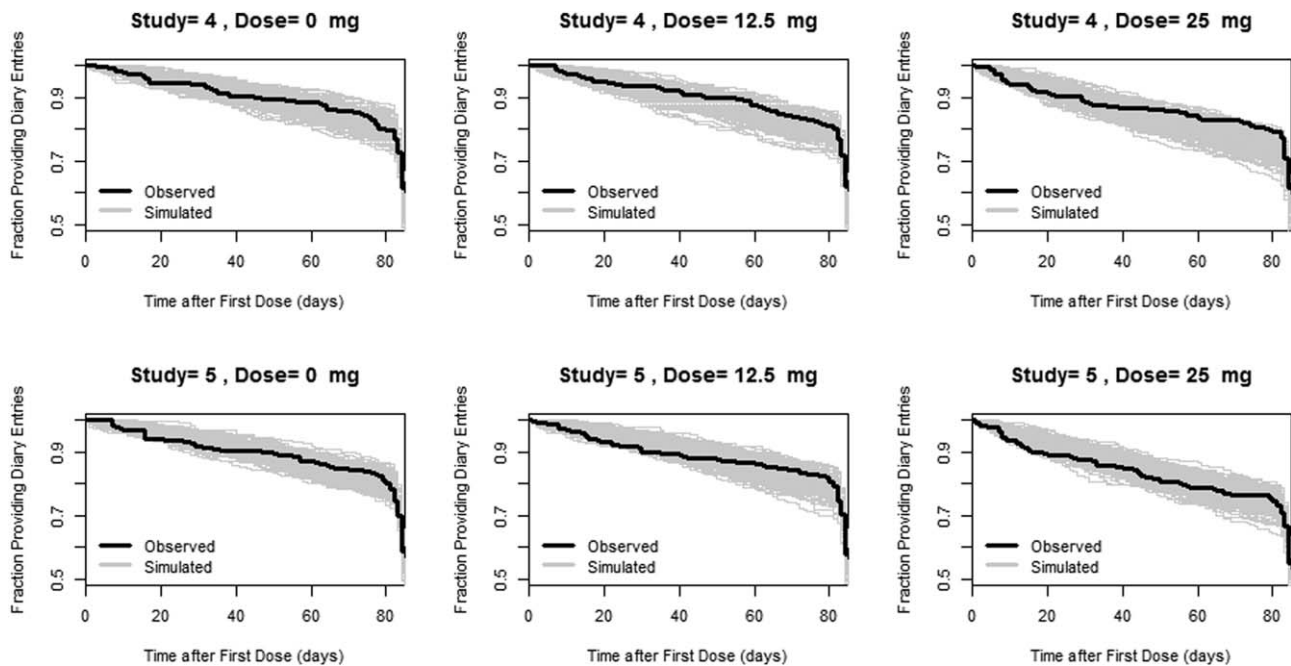
Figure 4 Kaplan–Meier visual predictive check for diary entry discontinuation model for study 4 (KODIAC 4) and study 5 (KODIAC 5).

number of SBMs was not determined to be an important predictor of DED and was not included in the base structural model.

The effects of age, sex, race, region, baseline opioid dose, and potency (weak vs. strong), and baseline laxative response were assessed simultaneously on the baseline log-hazard as well as the slope describing the change in the log-hazard with time over the double-blind treatment period. Covariates were not evaluated on the model component describing the sharp increase in DED near the end of the study. The full model included 22 additional covariate parameters and resulted in a 34.7 unit reduction in the OFV relative to the base structural model. For practical purposes, two covariate parameters with small estimated effect size magnitudes were excluded from the full model to derive the working full model in order to reduce the number of models that would subsequently need to be assessed with the WAM procedure (2^{20} vs. 2^{22} models). The working full model OFV differed from the full model by only 3.1 units. The WAM procedure identified the base structural model as the most parsimonious model indicating that none of the evaluated covariates were important predictors of DED. The rank correlation between WAM- and NONMEM-based SBC was 0.98, suggesting that the WAM procedure performed well. The final model predicted that patients in the naloxegol 25 mg/day group would have a slightly higher risk of early DED compared with the naloxegol 12.5 mg/day and placebo groups. Predicted discontinuation rates (90% CI) were 41.9% (37.5–46.3%) for naloxegol 12.5 mg/day and 43.4% (39.6–47.6%) for naloxegol 25 mg/day. The equation and parameter estimates for the final model are provided in **Supplementary Table S6**.

Table 2 Observed and population mean predicted responder rates

Treatment group	Observed responder rate, % (90% CI)	Population mean predicted responder rate, % (90% PI)
Placebo	29.5 (25.9–33.1)	32.9 (28.7–37.5)
Naloxegol 12.5 mg/day	38.1 (34.3–41.9)	42.7 (36.5–47.2)
Naloxegol 25 mg/day	41.9 (38.1–45.8)	43.0 (36.7–46.9)
By baseline opioid potency		
Placebo		
Weak	30.1 (24.1–36.2)	33.0 (26.3–39.7)
Strong	29.2 (24.4–34.0)	32.9 (27.6–38.3)
Weak + strong	28.9 (17.8–40.0)	32.6 (20.0–44.4)
Naloxegol 12.5 mg/day		
Weak	34.3 (27.7–40.9)	36.5 (27.9–44.3)
Strong	39.4 (34.4–44.4)	45.6 (37.5–51.0)
Weak + strong	42.9 (30.3–55.4)	45.9 (31.0–57.1)
Naloxegol 25 mg/day		
Weak	38.0 (31.1–44.8)	35.8 (27.7–43.1)
Strong	42.9 (37.8–48.0)	46.2 (38.6–51.4)
Weak + strong	47.3 (36.2–58.4)	46.6 (32.7–56.4)

The spontaneous bowel movement (SBM) model was used to simulate daily SBM counts in 500 datasets. The diary entry discontinuation (DED) model was then applied to each simulated dataset to replicate the DED rate from the observed data. Responder rate was calculated for each simulated dataset by treatment group. CI, confidence interval; PI, prediction interval.

Table 3 Population simulations of responder rates and diary entry discontinuation rates (DEDs)

Treatment group	Population mean (90% CI)		
	Responder rate	Placebo-adjusted responder rate	Diary entry discontinuation rate
Placebo	33.2 (30.2–36.0)	NA	42.0 (37.9–46.4)
Naloxegol 12.5 mg/day	43.5 (38.4–45.8)	10.4 (4.6–13.4)	41.9 (37.5–46.3)
Naloxegol 25 mg/day	44.2 (38.7–46.3)	11.1 (4.8–14.4)	43.4 (39.6–47.6)

Five-hundred sets of spontaneous bowel movement and DED model parameters were sampled using a smoothed bootstrap procedure. Each set of parameter estimates was used to simulate a dataset with 2,000 patients in each treatment arm. CI, confidence interval; NA, not applicable.

Visual predictive checks showed that the observed KM curves were contained within the distribution of simulated KM curves for each condition (**Figure 4**). Visual predictive checks using the binned hazard method[12] demonstrated that the model adequately captures the increase in risk of DED near the end of the study, since the observed empirical hazard rate was generally contained within the visual predictive check interval (**Supplementary Figure S2**). Plots limiting the magnitude of the y-axis to allow better resolution for early time bins also showed that the observed empirical hazard was mostly contained within the visual predictive check interval (**Supplementary Figure S3**). Overall, the visual predictive checks confirm that simulated DED data were consistent with observed DED data.

Predictive performance and population simulations for responder rate

Posterior predictive checks indicated that responder rates predicted by the models were in accordance with those observed during the clinical studies. The observed responder rates were contained within the 90% prediction intervals of the predicted responder rates for all treatment groups and for all baseline opioid potencies (**Table 2**). Simulated population mean responder rates were similar for the naloxegol 12.5 and 25 mg/day groups (43.5% vs. 44.2%; **Table 3**).

DISCUSSION

This population analysis explored the relationship between naloxegol exposure and response in patients with noncancer-related pain and OIC through the development of two models describing the number of SBMs per day and the time to DED. The SBM and DED models were combined to predict responder rates. An ordinal mixed-effects logistic regression dose–response model was developed to describe the daily number of SBMs. The logistic-regression approach was used to correct the lack of fit observed with negative binomial and Poisson models. DED rate was described by a time-to-event model. It was necessary to model DED rather than study or treatment discontinuation because all diary entries over the 12-week treatment period were used for calculation of the responder rate, regardless of when a patient discontinued

study medication or completed their final visit. Patients with fewer SBMs at baseline were predicted to have a smaller overall placebo effect and greater drug effect compared with patients with a higher baseline SBM frequency, suggesting that those with a higher pretreatment degree of constipation may gain the most benefit from naloxegol therapy.

As noted in the Methods, increased opioid potency was synonymous with increased opioid dose. The only demographic or baseline characteristic that had a significant effect on the relationship between naloxegol and daily number of SBMs was baseline opioid potency. Approximately 80% of patients were receiving low opioid doses at baseline ($<$200 morphine equivalent units (meu)/day) and $\sim$67% of patients were taking strong maintenance opioids ($\geq$200 meu/day) at baseline in studies 4 and 5. The median daily opioid dose was 132.5 meu/day in patients taking strong opioids compared with 45.0 meu/day in patients taking weak opioids. Since opioid potency is highly correlated with dose, it is difficult to assess the impact of opioid potency relative to dose. Therefore, it is difficult to make any conclusions regarding why the analysis favored potency over dose.

Patients taking strong opioids were predicted to benefit more from naloxegol therapy than those only using weak opioids. None of the demographic or baseline characteristics tested had a significant effect on DED.

Increased exposure to naloxegol was related to increased response, described in terms of a higher number of SBMs per day. However, the higher dose of naloxegol was associated with a slightly higher rate of early discontinuation compared with the lower dose. Thus, population mean responder rates computed by simulation were similar for the naloxegol 12.5 and 25 mg/day groups, since the increased frequency of SBMs in the 25 mg/day group was offset by a higher early discontinuation rate. The difference in observed responder rate between the 12.5 and 25 mg/day groups in the phase III clinical studies (38.1% vs. 41.9%) was larger than predicted (43.5% vs. 44.2%) but was within the expected variability anticipated for the observed sample sizes. Population predictions of placebo-adjusted responder rates (90% CI) were 10.4% (4.6–13.4%) and 11.1% (4.8–14.4%) for naloxegol 12.5 and 25 mg/day, respectively. Based on this exposure–response analysis, the 12.5 mg dose thus provided a clinical benefit over placebo with comparable efficacy to the 25 mg dose. However, naloxegol appeared to be more effective in patients with low baseline SBMs, suggesting that the 25 mg dose may provide additional benefit in patients with severe disease. Since the mean number of SBMs per week increased with increasing naloxegol dose and prespecified statistical significance was achieved for the primary endpoint in both pivotal trials for the 25 mg dose but in only one trial for the 12.5 mg dose, 25 mg was chosen as the starting dose.

In conclusion, the SBM and DED models could predict responder rates to naloxegol in patients with noncancer-related pain and OIC. The analyses support the conclusion that naloxegol provides therapeutic benefit, patients with more severe disease, and using strong opioids having the greatest response.

1. Panchal, S.J., Muller-Schwefe, P. & Wurzelmann, J.I. Opioid-induced bowel dysfunction: prevalence, pathophysiology and burden. *Int. J. Clin. Pract.* **61**: 1181–1187 (2007).
2. Bell, T.J. *et al.* The prevalence, severity, and impact of opioid-induced bowel dysfunction: results of a US and European Patient Survey (PROBE 1). *Pain Med.* **10**: 35–42 (2009).
3. Diego, L., Atayee, R., Helmons, P., Hsiao, G. & von Gunten, C.F. Novel opioid antagonists for opioid-induced bowel dysfunction. *Expert Opin. Investig. Drugs* **20**: 1047–1056 (2011).
4. Webster, L. *et al.* A phase 2, double-blind, randomized, placebo-controlled, dose-escalation study to evaluate the efficacy, safety, and tolerability of naloxegol in patients with opioid-induced constipation. *Pain* **154**: 1542–1550 (2013).
5. Pappagallo, M. Incidence, prevalence, and management of opioid bowel dysfunction. *Am. J. Surg.* **182**: 11S–18S (2001).
6. Chey, W.D. *et al.* Naloxegol for opioid-induced constipation in patients with noncancer pain. *N. Engl. J. Med.* **370**: 2387–2396 (2014).
7. Beal, S.L. & Sheiner, L.B. NONMEM users guides. NONMEM Project Group, University of California: San Francisco, CA, 1992.
8. Al-Huniti, N., Chapel, S., Xu, H., Bui, K.H. & Sostek, M. Population pharmacokinetics of naloxegol in a population of 1247 healthy subjects and patients. *Br. J. Clin. Pharmacol.* **81**: 89–100 (2016).
9. Kowalski, K.G. & Hutmacher, M.M. Efficient screening of covariates in population models using Wald's approximation to the likelihood ratio test. *J. Pharmacokinet. Pharmacodyn.* **28**: 253–275 (2001).
10. Yano, Y., Beal, S.L. & Sheiner, L.B. Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check. *J. Pharmacokinet. Pharmacodyn.* **28**: 171–192 (2001).
11. Huh, Y. & Hutmacher, M.M. Application of a hazard-based visual predictive check to evaluate parametric hazard models. *J. Pharmacokinet. Pharmacodyn.* **43**: 57–71 (2016).
12. Hutmacher, M.M. A visual predictive check for the evaluation of the hazard function in time-to-event analyses. Abstr 2940 $<$www.page-meeting.org/?abstract=2940$>$ (2013).