RESEARCH PAPER

Prediction of Disease Progression, Treatment Response and Dropout in Chronic Obstructive Pulmonary Disease (COPD)

F. T. Musuamba • D. Teutonico • H. J. Maas • A. Facius • S. Yang • M. Danhof • O. Della Pasqua

ABSTRACT

Purpose Drug development in chronic obstructive pulmonary disease (COPD) has been characterised by unacceptably high failure rates. In addition to the poor sensitivity in forced expiratory volume in one second (FEV_1), numerous causes are known to contribute to this phenomenon, which can be clustered into drug-, disease- and design-related factors. Here we present a model-based approach to describe disease progression, treatment response and dropout in clinical trials with COPD patients.

Methods Data from six phase II trials lasting up to 6 months were used. Disease progression (trough FEV_1 measurements) was modelled by a time–varying function, whilst the treatment effect was described by an indirect response model. A time-to-event model was used for dropout

Results All relevant parameters were characterised with acceptable precision. Two parameters were necessary to model the dropout patterns, which was found to be partly linked to the treatment failure. Disease severity at baseline, previous use of corticosteroids, gender and height were significant covariates on disease baseline whereas disease severity and reversibility to salbutamol/salmeterol were significant covariates on E_{max} for salmeterol active arm.

Conclusion Incorporation of the various interacting factors into a single model will offer the basis for patient enrichment and improved dose rationale in COPD.

KEY WORDS chronic obstructive pulmonary disease · disease modelling · disease progression · dropout · KPD model

F. T. Musuamba · D. Teutonico · M. Danhof · O. Della Pasqua (✉)
Gorlaeus Laboratories, Division of Pharmacology
Leiden Academic Centre for Drug Research
P.O. Box 9502, 2300 RA Leiden, The Netherlands
e-mail: odp72514@gsk.com

H. J. Maas · S. Yang · O. Della Pasqua
Clinical Pharmacology Modelling & Simulation, GlaxoSmithKline
Uxbridge, UK

A. Facius
Department of Pharmacometrics, Nycomed GmbH
Konstanz, Germany

ABBREVIATIONS

BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
CTS	Clinical trial simulations
EDK_{50}	Apparent potency parameter corresponding to the exposure associated with half of the maximum effect
Emax	Maximum effect
FEV_1	Forced expiratory volume in one second
FOCEI	First order conditional estimation method with interaction
GOLD	Global Initiative for Chronic Obstructive Lung Disease
Int_Dis	The disease status at baseline, *i.e.*, at the start of the clinical trial
Kin	Zero-order process parameterised as a synthesis rate constant
Kout	First-order elimination constant
KPD	Kinetic-pharmacodynamic
MOFV	Minimum value of objective function
PICS	Previous use of corticosteroids
PsN	Perl-speaks-NONMEM
Slope_Dis	Parameter describing the daily decline of FEV_1 due to disease progression
VPC	Visual predictive check

Springer

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a chronic respiratory disorder that progresses slowly and is characterised by an obstructive ventilatory pattern, which is rarely reversible, very often related to tobacco smoking and which can lead to chronic respiratory failure (1). COPD is a major public health problem. Indeed, it is the fourth leading cause of morbidity and mortality in industrialised countries and is projected to rank fifth in 2020 in burden of disease caused worldwide by the World Health Organization (2,3). Even though exacerbations have been recognised as an important trait of the disease and is increasingly included in the evaluation of drug effects, forced expiratory volume in one second (FEV_1) remains the most frequently used endpoint in clinical trials in COPD, with change from baseline (ΔFEV_1) corrected for placebo being the primary measure of efficacy. In fact, FEV_1 is currently considered to be one of the best predictors of patient survival and as such has been used as a prognostic marker for outcome (4). Consequently, regulatory authorities still recommend the use of measures of airflow obstruction improvement for registration trials of new bronchodilators.

From a therapeutic perspective, maintenance treatments are based on the use of bronchodilators which rapidly improve lung function. However, high failure rate is observed during the development of treatments for COPD. Numerous factors are known to contribute to this phenomenon, which can be clustered into drug-, disease- and design-related factors. Of particular relevance are the sensitivity of spirometric measures to patient demographic characteristics and the disease-related inclusion criteria such as gender, previous use of inhaled corticosteroids and disease severity at baseline. Yet, the majority of trials are usually conducted according to a fixed protocol, copied from one study to the next. Among the reasons for this conservatism are the high costs of these trials, which limit the willingness to experiment with designs, and concerns about potential failure because of deviations from traditional study designs. Another important point supporting current practice is the statistical evaluation of efficacy (*i.e.*, contrast between active agent and placebo) by changes from baseline at a single time point.

In other therapeutic areas in which similar concerns exist regarding the impact of confounding design factors, costs and time course of the response, the use of modelling and simulation has been advocated. In fact, the implementation of clinical trial simulations (CTS) allows evaluation of a single factor at a time. In addition, it provides the opportunity to investigate factors that have not been implemented yet and can therefore not be included in a meta-analysis. CTS enable statistical inferences from models and as such contrasts with the long-established calculations of statistical power, which are often based only on a point estimate of the variability of the clinical end point.

Here we show how a model-based approach can be used to discriminate drug effects from disease- and design-related factors on the overall treatment response. The aim of this investigation was therefore (1) to develop and validate a model for trough FEV_1 including disease progression and dropout; and (2) to quantify the impact of relevant demographic and disease-related factors on treatment response. Our analysis is based on a kinetic-pharmacodynamic (KPD) model which was previously developed to overcome some of the limitations associated with lack of pharmacokinetic data. (5,6) This is a common problem during the development of inhaled compounds, for which plasma pharmacokinetics is not available and/or doses do not relate directly to treatment response. This concept has been successfully applied to model the effects of a novel long-acting bronchodilator in COPD in a recently published study (7).

Given the typical duration of efficacy trials in COPD, disease progression and dropout models were included to ensure unbiased estimation of treatment response. Dropout is considered to be informative (also called non random) when it is driven by the unobserved clinical response and non-informative otherwise (also called random and completely random dropout). In contrast to non-informative dropout that can be just described or predicted independently to the clinical response of interest, informative dropout must be jointly modelled with the response data in order to warrant reliable parameter estimates. Further characterisation of the patterns of dropout is therefore essential for models aimed at the prediction and simulation of response across a wide time span. Joint analysis of disease progression, treatment response and dropout is very often the most objective and reliable way to assess dropout patterns and account for its effects on parameter estimation. Joint models have been developed to describe the time course of biomarkers in long term diseases associated with relatively high dropout rates such as Parkinson's disease (7,8), depression (9) and chronic viral infections (10). To date, joint modelling concepts have not been applied to FEV_1 response in COPD.

METHODS

Patients

The clinical data used for the analysis were extracted from Top Institute Pharma data repository. The data were anonymised and randomly combined prior to use. It consisted in a pool from 6 Phase II multicenter studies, containing placebo ($n = 2273$) and active treatment (salmeterol) ($n = 270$). All patients had a diagnosis of moderate to severe COPD as defined by the Global initiative for Chronic Obstructive Lung Disease (GOLD) guidelines (11). Moderate COPD implies that post-bronchodilator FEV_1 values at the start of the study

(T_0) are between 50 and 79% of the predicted FEV_1 for a subject with normal lung function, whereas for severe COPD post-bronchodilator FEV_1 values range between 30 and 50% of the predicted FEV_1. Very severe patients (post-bronchodilator FEV_1 at $T_0 < 30\%$) were excluded from the trials. Spyrometry (trough FEV_1) was performed after administration of bronchodilators at intervals ranging between two and a maximum of four weeks by trained clinical staff. The study drugs were inhaled salmeterol (50 µg) or matched placebo administered as a twice daily regimen. Patients were not allowed to take another bronchodilator during the course of treatment, other than rescue medication (e.g., in case of exacerbations). Table I summarises patient demographics and disease related factors at the start of the trials, including study duration and number of patients enrolled. The protocols were approved by the local institutional review boards, and all patients agreed to sign an informed consent form. We have used data from patients enrolled into three of the six studies for model building purposes and included those in the remaining studies during the model validation procedures. The final model parameter estimates refer therefore to the total pool of patients in all six studies.

Modelling and Covariate Analysis

Nonlinear mixed effects modelling was performed using NONMEM v7.1.2 (double precision, Icon Development Solutions, Ellicott City, MD, USA) and Perl-speaks-NONMEM (PsN)-toolkit, (12) a programming library containing a collection of computer intensive statistical methods for non-linear mixed effects modelling, Xpose 4.0 (13) and R (14). First order conditional estimation method with interaction (FOCEI) was first used to model disease progression and drug effects. Subsequently, dropout patterns were analysed using the Laplace estimation method, which is deemed to be more appropriate to jointly model continuous and categorical data in NONMEM. The models for disease progression, drug effects and dropout were developed simultaneously. For the sake of completeness, the disease and drug effect parameter estimates obtained with and without dropout were compared to assess the influence of informative dropout.

Base Model

Disease and Drug Effects Model. The process describing FEV_1 changes over time was modelled using an indirect response model, as described by the following differential equation

$$\frac{dFEV_1}{dt} = Kin - Kout \times FEV_1 \tag{1}$$

Table I COPD Patient Demographics and Study Characteristics

	Median [range]
Number of patients (studies 1–6)	2543
Reversibility (mL)	56
Age (years)	65 [40–90]
Sex (% of females)	30
Height (cm)	170 [135–203]
Patients with severe COPD (%)	67
Patients with reversible COPD (%)	26
Patients with a previous use of inhaled corticosteroids (PICS) (%)	40
Smoking status (% of smokers)	42
Dropout (% of dropout)	13
Study duration (weeks)	167

where the change in the observed FEV_1 over time ($dFEV_1/dt$) is controlled by a zero-order process parameterised as a synthesis rate constant (Kin) and first order elimination process (Kout). The disease status at baseline was considered to be described by the ratio between Kin and Kout. The disease progression was modelled as a linear decline in the baseline status as follows:

$$Dis = Int\_Dis - Slope\_Dis \times Time \tag{2}$$

with

$$Dis = \frac{Kin}{Kout} \tag{3}$$

where Dis is the disease status, Int_Dis is the disease status at the start of the clinical trial, and Slope_Dis is the daily decline of Dis due to disease progression.

The effect of bronchodilators was modelled using a non-linear Emax function, in which the maximum effect is proportional to an apparent potency parameter (EDK_{50}). Four different models were tested with drug effects incorporated either additively or multiplicatively on Kin or on Kout. Examples are provided below for incorporation on Kin:

$$Kin = TVKin + \left(\frac{A \times Emax}{EDK_{50} + A} \right) \tag{4}$$

$$Kin = TVKin \times \left(1 + \frac{A \times Emax}{EDK_{50} + A} \right) \tag{5}$$

where TVKin is the typical value of Kin in the absence of drug effects, EDK_{50} is the drug exposure associated with half of the maximum effect, Emax is the maximum drug effect, and A is the drug exposure marker. Different parameterisations were tested for characterisation of the exposure markers including:

a. $d(A)/dt = -KDE \times A$ and $A_0 = Dose$
b. $A = Dose$
c. $A = D \times KDE$ with $d(D)/dt = -KDE \times D$ and $D_0 = Dose$

where A is the amount (a), the dose (b), or the input rate (c) of drug into the effect compartment, D is the dose at time 0 KDE is the first order elimination rate constant for the bronchodilator and A_0/D_0 are the initial values for A/D.

Given that all the patients belonging to the active arms of the different trials were taking the same dose of bronchodilators, a prior function ($PRIOR) was used in NONMEM to stabilise the estimation of correlated parameters (*i.e.*, Emax and EDK_{50}). In addition, the ratio between Emax and EDK_{50} was first estimated and EDK_{50} was computed subsequently. Prior information on the exposure response relationship was derived from a model using data from a previously published study (15). Emax and EDK_{50} distributions were directly taken from the NONMEM output file of the previous model. The parameter estimates for Emax and EDK_{50} were 0.78 and 4 μg, respectively. The values used as priors on interindividual variability (OMEGA) were 0.24 and 0.14 with a covariance term of 0.1. The NONMEM variance/covariance matrix was used to weight the prior information. The degrees of freedom for the OMEGAs were based on the number of patients included in the historical study and set to 24.

Dropout model. Dropout was modelled using a time-to-event model (10,16,17). The probability of a having a dropout at any given time was predicted by describing the hazard (h(t)) associated with the dropout (*i.e.*, event). Different hazard models were tested including constant, exponential and Weibull models, as previously described in (16,17).

As time passes the cumulative hazard predicts the risk of having the event over the interval 0-t. The risk (cumulative hazard) was obtained by integrating hazard with respect to time:

$$Cumulative\ hazard = \int_0^t h(t) \tag{6}$$

On the other hand, the probability of remaining in the trial (*i.e.*, of not dropping out) was predicted by a survivor function from the cumulative hazard.

$$Survivor(t) = P(T > t) = \exp\left(-\int_0^t h(t)\right) \tag{7}$$

The probability of dropping out at any given time was predicted by the probability density function (pdf(t)). The pdf

was calculated from the survivor function and from the hazard at the specific time as follows:

$$pdf(t) = Survivor(t) \times h(t) \tag{8}$$

The probability of dropping out within an interval was modelled by taking the difference between the survivor at the first time after dropout and the survivor at last observed time before dropout.

Dropout patterns were characterised using likelihood ratio tests by including additional parameters on the initial hazard term. For example when a constant hazard model was tested, the following equation was used:

$$h(t) = \beta_0 \times e^{\beta_1 \times FEV_1 + \beta_2 \times IPRED} \tag{9}$$

where IPRED represents the individual predicted FEV_1 at time (t), FEV_1 is the observed measure at available sampling times, the coefficient β_0 and the exponents β_1 and β_2 describe missingness completely at random, at random and non-random, respectively.

Inter-Individual and Unexplained Variability Model. Exponential models were used for inter-individual variability (IIV) on model parameters. The value of a parameter in the ith individual (P_{ij}) was a function of the typical value of the parameter (TVP) and of the individual deviation initially represented by η_i, which describes the inter-patient variability term for the ith patient. The ηs in the population were assumed to be normally distributed random variables with zero mean and a variance that is estimated as part of the model:

$$P_{ij} = TVP \times \exp(\eta i) \tag{10}$$

The covariance between different parameters was also assessed using additional random effects terms using the OMEGA BLOCK() option in NONMEM. η terms were only maintained in the model when they improved the fitting as assessed by comparing the NONMEM objective function values with and without the inclusion of the random effects. A difference of at least 3.84 points was necessary to keep the random effects term in the model.

Additive, proportional and mixed error models were tested for the residual error as shown in Eqs. 11–13:

$$Y = IPRED + \varepsilon_{add} \tag{11}$$

$$Y = IPRED \times \left(1 + \varepsilon_{prop}\right) \qquad (12)$$

$$Y = IPRED \times \left(1 + \varepsilon_{prop}\right) + \varepsilon_{add} \qquad (13)$$

where Y and IPRED represent the observed and individual predicted FEV_1 values, respectively. ε_{add} and ε_{prop} are the additive and the proportional error terms on FEV_1 values, respectively. Epsilon was supposed to be normally distributed, zero-mean random variables with variance (σ^2) terms that are estimated as part of the population model-fitting process.

Model Selection. Structural model selection was based on 1) the change in the minimum objective function value (MOFV), with a decrease of at least 3.84 points being required for additional parameter inclusion, 2) the plausibility and the precision of parameter estimates and 3) the qualitative evaluation of goodness-of-fit plots. The precision of parameter estimates, expressed as standard error of estimates, was generated by the covariance option in NONMEM. Goodness-of-fit plots comprised predicted and individually predicted *versus* observed data, conditional weighted residuals (18) (CWRES) *versus* time after dose, as well as the normalised prediction distribution errors (NPDE) (19). The precision of parameter estimates, expressed as standard error of estimates, was generated by the covariance option in NONMEM.

Covariate Model

A stepwise inclusion/exclusion approach was used to investigate the relationship between different structural model parameters and demographic factors. To explain inter-patient variability on disease and on drug-related parameters, and to reduce the unexplained residual variability, relationships were investigated for age, sex, height, BMI, weight, severity, reversibility, previous use of inhaled corticosteroids (PICS). Individual empirical Bayes estimates of parameters were generated and their correlation with each covariate being evaluated separately during the exploratory steps of the covariate analysis. The log-likelihood ratio is assumed to be χ^2-distributed. A difference in objective function value of 3.84 is considered to be significant at $p < 0.05$ with one degree of freedom (difference of one parameter between the two nested models). Continuous variables (age, weight, height) were centred to their median values and tested on the parameters using a power model (see Eq. 14). Categorical variables were entered in the model in a multiplicative manner (see Eq. 15). This is exemplified below for height and gender:

$$Int\_Dis = TVInt\_Dis \times \left(\frac{HT}{HTmed}\right)^{\theta_{cov}} \qquad (14)$$

$$Int\_Dis = TVInt\_Dis \times \theta_{cov}[in\ case\ of\ males] \qquad (15)$$

where Int_Dis is the disease baseline status (*i.e.*, baseline FEV_1 at $t=0$ expressed in L), TVInt_Dis is the typical value for disease baseline intercept (in case of continuous covariate, TVInt will be the disease baseline status for a patient with median covariate (here height) value and for categorical covariates, it will be the value for the reference category (here females)), Θ_{COV} is the relative change in baseline intercept due to covariate effect. HT is the patient body height and HTmed is the median body height in the dataset.

A backward deletion process was used to complete the covariate model building. Only covariates which upon deletion caused an increase in MOFV>11 points (χ^2 p-value≤ 0.001) were retained in the final model.

Final Model Evaluation

An initial assessment of the predictive performance of the model was carried out by using the parameter estimates from the model building studies 1–3 ($n=1154$ patients) to describe the trough FEV_1 response profile in the remaining studies 4–6 ($n=1389$ patients). Bootstrapping and simulations were used to internally validate the final model. 100 bootstraps were generated using PsN and confidence intervals were obtained for all model parameters. NPDE-related diagnostic plots (19) and mirror plots were then generated to assess overall model performance.

Finally, visual predictive checks (VPC) were performed taking dropout into account. Observed and simulated FEV_1 values *vs.* time were presented in conjunction with Kaplan Meier plots to assess the suitability of the dropout model.

RESULTS

Disease progression (trough FEV_1 measurements) was modelled by a time–varying function. The treatment effect was described by an indirect response model using an Emax model in a multiplicative manner on Kin. In addition, a time-to-event model with constant hazard was used for dropout. A schematic representation of the structural model for disease progression and drug effects is shown in Fig. 1. The following disease progression and drug-related parameters were estimated: disease baseline intercept (Int_Dis), disease baseline slope (Slope_Dis), zero order input rate constant for the biological response (Kin), first order elimination rate constant (KDE), maximum drug effect (Emax) and the ratio between the apparent potency (EDK_{50}) and Emax. Two parameters (β_0 and β_2) were necessary to model the dropout patterns. A constant hazard (β_0) of 0.06 was estimated, which was

Fig. 1 Schematic representation of the structural model. Bronchodilatory effect on FEV_1 is described by an Emax function using a KPD model as input for drug exposure. The overall response to treatment accounts for the course of disease, *i.e.*, the natural decrease in FEV_1 over time, which has been parameterised in terms of an indirect response model. See text for further explanation of the parameters.

$$Effect = 1 + \frac{A \times Emax}{EC_{50} + A}$$

$$with \quad Kout = \frac{Kin}{Int\_Dis - slope\_Dis \times Time}$$
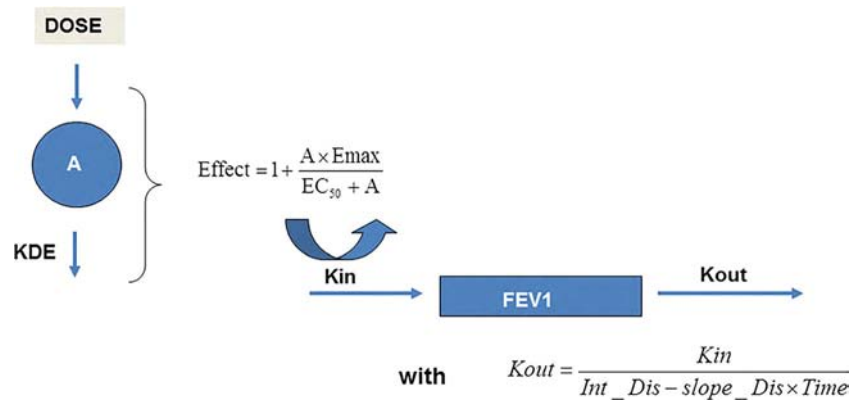
increased in case of treatment failure (*i.e.*, lower FEV_1 values). The inclusion of β_2 in the model led to a decrease of 15 points in the MOFV as compared to the model only including β_0. The best fit was obtained with interindividual variability included on the disease slope, the disease baseline intercept, Emax, KDE, EDK_{50} and Kin. All covariance terms turned to be lower than 0.1 and their removal did not appear to decrease model performance or goodness of fit. Random effects were not imputed on dropout parameters. Residual errors were best described by an additive model. The parameter estimates for the base model are summarised in Table II.

During covariate model building, disease severity at baseline, previous use of inhaled corticosteroids, gender and height were retained as significant covariates on Int_Dis whereas reversibility to salbutamol/salmeterol and disease severity at baseline were significant covariates on Emax for the active arm (salmeterol). All covariates were included in the final model in a as shown in Eqs. 14 and 15. A summary of the final parameter estimates including the aforementioned covariates is provided in Table II. All model parameters were well estimated, as shown by the relative low standard error of estimates (<50%) and by shrinkage levels <30% for all random effects parameter values. Basic goodness-of-fit plots also showed good model predictive performance (see Fig. 2).

Validation of the final model by external validation procedures and bootstrapping yielded good predictive

Table II Overview of Base and Final Model Parameter Estimates

Parameter	Base model [RSE][a]	Final model [Bootstrap 90% CI]
Θ_{TVEmax}	0.34 [0.04;11.5]	0.36 [0.05;0.81]
$\Theta_{EDK50/Emax}$[b] (μg)	10.33 [0.64;20.02]	8.66 [5.19;11.39]
Θ_{KDE} (hr^{-1})	54.46 [1.64;1806]	52.46 [12.18;65.14]
Θ_{Int_Dis} (disease intercept) (L)	0.9 [0.3;2.99]	1.16 [1.06;1.19]
Θ_{Slope_Dis} (disease slope) (L/day)	−0.82 [−1.93;-0.02]	−0.596 [−0.214;]
Θ_{Kin} (L/day)	0.04 [0.001;1.32]	0.0004 [0.0001;0.0032]
$\Theta_{severity\ on\ Emax}$	–	0.94 [0.08;0.99]
$\Theta_{reversibility\ on\ Emax}$	–	1.05 [1.02;1.52]
$\Theta_{severity\ on\ Int\_Dis}$	–	0.62 [0.22;0.84]
$\Theta_{sex\ on\ Int\_Dis}$	–	0.82 [0.13;0.97]
$\Theta_{PICS\ on\ Int\_Dis}$	–	0.92 [0.77;0.94]
$\Theta_{height\ on\ Int\_Dis}$	–	1.90 [1.48;2.52]
Initial hazard term (β_0)	0.006 [0.0001;0.33]	0.006 [0.000;0.041]
Proportional hazard term for informative dropout (β_2)	−0.843 [−0.11;−1.57]	−0.880 [−0.19;−0.92]
IIV_{EDK50} (CV%)	8 [0.01;15.9]	–
IIV_{Emax} (CV%)	76 [0;153]	70 [32;108]
IIV_{Int_Dis} (CV%)	35 [0;81]	25 [12;63]
IIV_{Slope_Dis} (CV%)	16 [0;33]	–
IIV_{Kin} (CV%)	10 [3;19.7]	8 [11;28]
σ (L)	0.36 [0.021;0.39]	0.13 [0.08;0.24]

[a] *RSE* relative standard errors; values computed from standard error of estimates generated by covariance step in NONMEM

[b] EDK_{50} estimates were based on a parameterisation relative to Emax
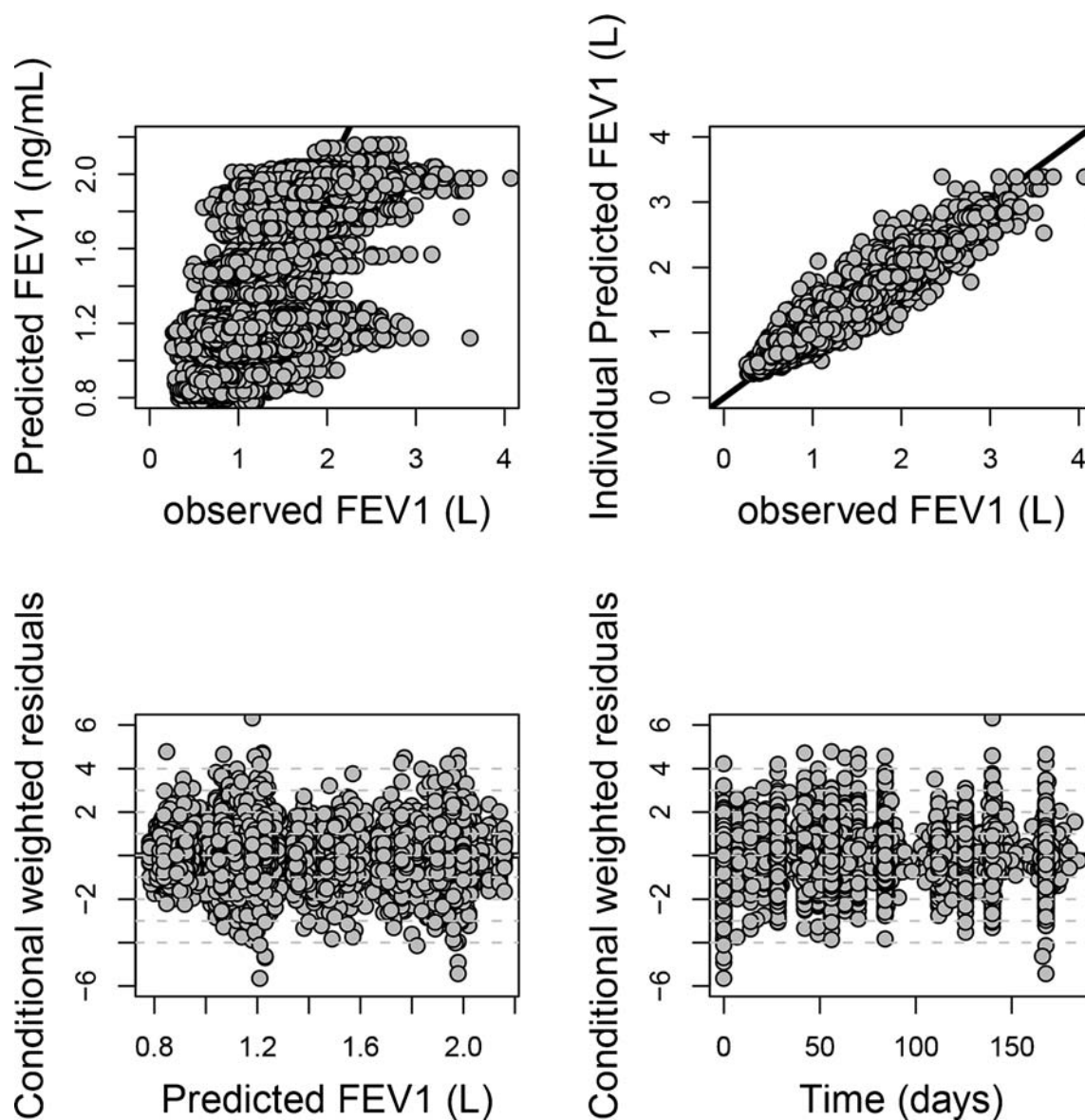
Fig. 2 Basic goodness-of-fit plots for the final model, including population and individual predicted *vs.* observed FEV$_1$, conditional weighted residuals *vs.* time and individual predicted FEV$_1$.

performance. All observed parameter values obtained by bootstrapping were found to be within the 90%-confidence interval ($n = 100$ bootstraps). In addition, as shown by the visual predictive check in Fig. 3, most of the observed data in the validation subset of data were distributed within the 5th and 95th percentiles of the prediction intervals. For the sake of clarity, similar predictive performance is observed when the response profiles were split by severity status (severity status being the covariate with the highest effect size) and by dose group. The overlap between predicted and original distributions shows that the model accurately captures drug, disease and dropout effects (Fig. 4).

One of the interesting findings from our analysis is that the change in FEV$_1$ values over time was found to be dependent on the baseline characteristics. For instance, as shown in Fig. 5, patients with a more severe disease status and unfavourable covariate characteristics at baseline will have a slower evolution over time and will be less sensitive to drug effects (see Fig. 5(a) and (b)). The correlation between change in FEV$_1$ values over time and baseline characteristics is consistent with indirect drug effects, as described by increasing zero-order processes. This is particularly relevant for salmeterol, given that baseline severity status shows a higher effect on the disease slope than the active treatment itself when salmeterol is administered at 50 µg doses according to a twice daily-dosing regimen. On the other hand, higher sensitivity to the drug effects was more pronounced in patients who show reversibility to salbutamol/salmeterol at baseline.
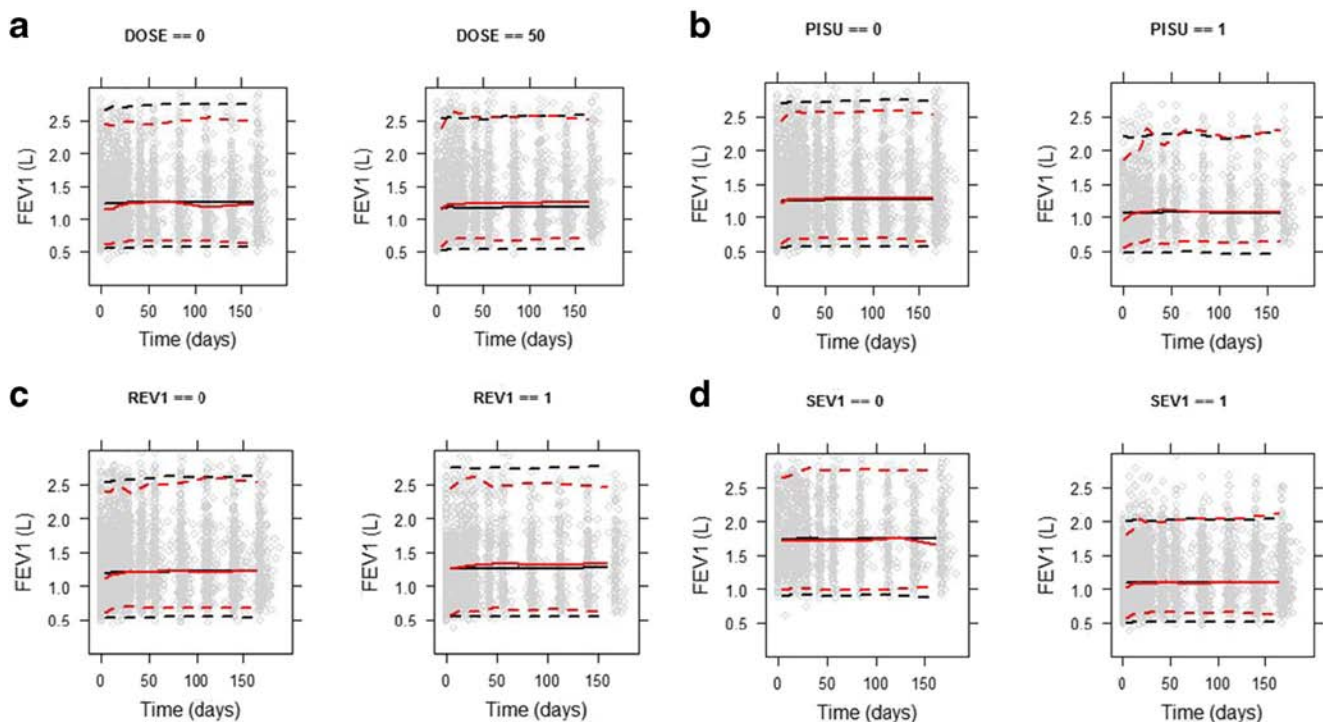
Fig. 3 Visual predictive check stratified by (**a**) treatment (*left panel*: placebo, *right panel*: 50 μg dose), (**b**) by previous use of inhaled corticosteroids [PISU] (*left panel*: no, *right panel*: yes), (**c**) by reversibility to salbutamol/salmeterol [REV] (*left panel*: non-reversible, *right panel*: reversible), and (**d**) by severity status [SEV] (*left panel*: non-severe, *right panel* severe). *Grey dots*: observed concentrations, *black dotted lines*: limits of the 95% prediction intervals for the observations, *black dotted lines*: limits of the 95% prediction intervals for the simulations, *black continuous line*: median line for the observations, *red continuous line*: median line for the simulations. See text for further details on the scenarios selected for the VPC.

DISCUSSION

We have shown the performance of a joint model for trough FEV_1 in a large COPD patient population, which takes into account disease progression, drug effects and dropout. Despite the wide variation in the individual time course profiles, our

Fig. 4 Kaplan Meier plots of observed (*black*) and simulated (*grey*) dropout for a subset of the pooled studies used for the current analysis.

approach resulted in a model with appropriate predictive performance, as shown by the precision in parameter estimates and the quality of goodness of fit for both treated and untreated arms of the study population.

Albeit not explicitly mechanistic, this model is consistent the known underlying processes associated with the pathophysiology of COPD. In fact, a more descriptive parameterisation of the underlying disease processes may not be feasible with data from typical Phase II clinical trials. Moreover, previous research has shown that no clear correlation exists between biomarkers of the underlying inflammatory condition and common spirometric measures, such as trough FEV_1 or serial FEV_1, making it difficult to define a fully mechanistic parameterisation for disease processes (20,21). Treatment response on FEV_1 was therefore best described by an Emax model.

Parameterisation of the disease processes was based on the assumption that spirometric measures result from a series of putative turnover (*i.e.*, synthesis and degradation) processes, which can be modelled as a zero-order input rate and leading to FEV_1 increase and a first-order elimination rate leading to FEV_1 decrease. The synthesis-related processes include all the factors that can enhance the respiratory capacity, such as the ciliary function, the physiological beta-adrenergic tonus and the drug-related effects (*i.e.*, beta agonism and muscarinic
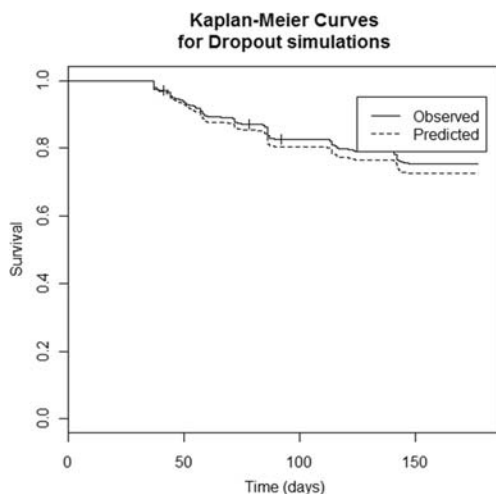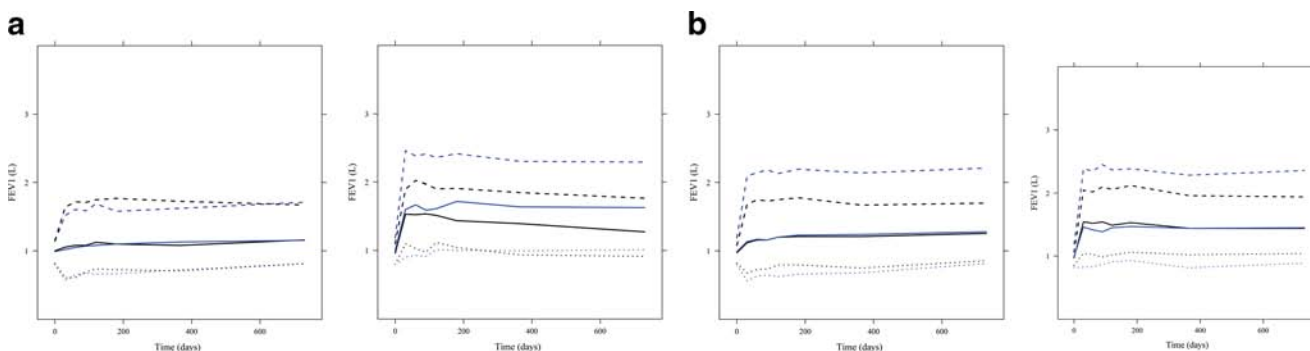
Fig. 5 Simulations ($n = 1000$) of four scenarios based on the final model. In *blue*: median and 95% prediction intervals for the placebo and in *black*: median and 95% prediction intervals for the active arm. (**a**) *left panel*: non-reversible and severe patients, right panel: reversible and non-severe patients; (**b**) patients with height >172 cm and previous use of inhaled corticosteroids, *right panel*: patients with height >172 cm and no previous use of inhaled corticosteroids. To ensure appropriate comparisons between the groups, the remaining covariates were kept in a balanced manner in each group.

antagonism) whereas the degradation-related processes comprise the secretory (mucus production) function of the epithelial cells and inflammatory components which restrict airway reversibility in COPD. This type of parameterisation has been used to describe several physiological responses involving inflammatory conditions (22–24). As indicated previously, our modelling approach incorporated disease progression, which is expressed as a time-varying increase in the ratio between degradation and synthesis of response. This model is therefore consistent with the results of several previously reported trials (26–30) showing that bronchodilators effects (including β_2-agonists, such as salmeterol and M_2 antagonists, such as tiotropium) in COPD are characterised by a rapid improvement of the lung function followed by a long steady phase. In addition, our analysis replicates the unexplained variability observed in FEV_1 in previously published models, which was estimated to be approximately 0.1 L. The additive residual variability in FEV_1 was 0.13 L in our model.

The different covariates included in our final model were measured or recorded at the start of the studies (disease severity, gender, previous use of inhaled corticosteroids, body height). Whilst their role in the long-term prognosis of COPD has been previously demonstrated (31,32), this is the first time that their influence on FEV_1 is described in a quantitative manner using a joint model, in which dropout and disease progression are considered concurrently. In addition, our choice of parameterisation has enabled the discrimination between disease and drug-related changes in response. This approach also shows how dropout affects drug-specific parameters. Indeed, when fitting the model to the data without taking dropout into account, differences of 37 and 55% where found in the obtained estimates of EDK_{50} and Emax, respectively.

Today's pursuit for disease-modifying drugs and personalised therapies requires one to understand the determinants of response and variability as the basis for subsequent evaluation of candidate molecules and tailored dosing

regimens. Equally important is our understanding of the role of such factors on trial design during the evaluation of safety and efficacy (33). Unfortunately, for many clinical experts modelling results do not provide immediate evidence of the impact that such a parametric representation of disease progression can have on the development of novel anti-inflammatory drugs for COPD. From a theoretical perspective, the availability of drug- and system-specific parameters enables us to predict the outcome of novel interventions and handle potential confounders of response as covariates. Furthermore, our analysis clearly shows that despite the controversy regarding the use of FEV_1 as a marker of disease progression (34), further insight into the underlying disease processes can provide the basis for patient selection and enrichment in clinical trials.

As previously stated, the primary purpose of this model will be its use in clinical trial simulations. The added value of scenario analysis and quantitative assessment of the impact of enrichment procedures on treatment response in Phase IIa and IIb trials will be the object of separate report in which a series of trial designs and treatment options are scrutinised. Yet, from a purely statistical perspective, two immediate lessons can be learnt from this analysis. First, the inclusion of a dropout model has highlighted the implication of non-random data missingness in longitudinal efficacy and safety trials. Of particular interest is the underestimation of the disease status (described by FEV_1 values) and consequently inaccurate estimation of the drug effects. Second, the potential bias in effect size when comparing changes in FEV_1 relative to baseline (ΔFEV_1) for different treatment arms (active arm *vs.* placebo) based on traditional hypothesis testing (*e.g.*, unpaired ANOVA test). This approach can easily fail to identify drug effects when confounding effects such as baseline characteristics are not explicitly incorporated as covariates.

We need to acknowledge a few limitations in the analysis described here. It must be noticed that parameter identifiability issues arise when sampling intervals are wide

and data on different treatment levels do not exist (*i.e.*, only one active dose level is usually evaluated in a clinical protocol). This issue results in parameter uncertainty, as shown by the wide confidence intervals obtained for most parameter estimates. This limitation is unlikely to be overcome unless phase II protocols are designed with the explicit objective of characterising exposure-response relationships (33). In addition, estimation of drug-specific parameters such as Emax and EC_{50} could not be obtained solely from the available clinical data. The use of prior information was required under the assumption of comparable, exchangeable experimental conditions across studies. We also emphasise the role of exclusion criteria on the identification of covariates and subsequent estimation of covariate effect size. Prediction of treatment effect in a real-life population requires one to account for those patients who are normally excluded from clinical trial protocols.

CONCLUSION

Using a joint model, we have shown that gender, height, disease severity at baseline, previous use of corticosteroids and dropout rates should be taken into consideration when designing early trials. We anticipate its value in the design of prospective trials for which currently used statistical approaches are inadequate or incomplete. Incorporation of the various interacting factors into a single model will offer the basis for patient enrichment and improved dose rationale in COPD.

ACKNOWLEDGMENTS AND DISCLOSURES

REFERENCES

1. MacNee W. Pathogenesis of chronic obstructive pulmonary disease. Proc Am Thorac Soc. 2005;2:258–66.

2. Raherison C, Girodet P-O. Epidemiology of COPD. Eur Respir Rev. 2009;18:213–21.

3. Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, Atkinson C, et al. UK health performance: findings of the Global Burden of Disease Study 2010. Lancet. 2013;381:997–1020.

4. Franciosi LG, Page CP, Celli BR, Cazzola M, Walker MJ, Danhof M, et al. Markers of exacerbation severity in chronic obstructive pulmonary disease. Respir Res. 2006;7(74):1–14.

5. Jacqmin P, Snoeck E, van Schaick EA, Gieschke R, Pillai P, Steimer JL, et al. Modeling response time profiles in the absence of drug concentrations; definition and performance evaluation of the K-PD model. J Pharmacokinet Pharmacodyn. 2009;34:57–85.

6. Holford NH, Peace KE. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. Proc Natl Acad Sci U S A. 1992;89:11466–70.

7. Wu K, Looby M, Goonaseelan Pillai G, Pinault G, Drollman AF, Pascoe S, et al. Population pharmacodynamic model of the longitudinal FEV1 response to an inhaled long-acting anti-muscarinic in COPD patients. J Pharmacokinet Pharmacodyn. 2011;38:105–19.

8. Kimko HC, Reele SS, Holford NH, Peck CC. Prediction of the outcome of a phase 3 clinical trial of an antischizophrenic agent (quetiapine fumarate) by simulation with a population pharmacokinetic and pharmacodynamic model. Clin Pharmacol Ther. 2000;68:568–77.

9. Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. J Biopharm Stat. 2001;11:9–21.

10. Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn. 2003;30:83–103.

11. American Thoracic Society. Standardization of spirometry, 1994 update. Am J Respir Crit Care Med. 1995;152:1107–36.

12. Lindbom L, Pihlgren P, Jonsson EN. PsN-Toolkit–a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. Comput Methods Prog Biomed. 2005;79:241–57.

13. Jonsson EN, Karlsson MO. Xpose-an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM. Comput Methods Prog Biomed. 1999;58:51–64.

14. R Development Core Team. R: a language and Environment for Statistical computing. In R Foundation for Statistical Computing (ed.), Vienna Austria, 2009.

15. Kemp JP, Bierman CW, Cocchetto DM. Dose–response study of inhaled salmeterol in asthmatic patients with 24-hour spirometry and Holter monitoring. Ann Allergy. 1993;70:316–22.

16. Rubin DB. Inference and missing data. Biometrika. 1976;63:581–92.

17. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. Appl Stat. 1994;43:49–93.

18. Hooker AC, Staatz CE, Karlsson MO. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. Pharm Res. 2007;24:2187–97.

19. Comets E, Brendel K, Mentre F. Computing normalized prediction distribution errors to evaluate nonlinear mixed effects models: the npde add-on package for R. Comput Methods Prog Biomed. 2008;90:154–66.

20. Gabrielsson J, Jusko W, Alari L. Modeling of dose–response-time data: four examples of estimating the turnover parameters and generating kinetic functions from the response profiles. Biopharm Drug Dispos. 2000;21:41–52.

21. Franciosi LG, Page CP, Celli BR, Cazzola M, Walker MJ, Danhof M, et al. Markers of disease severity in chronic obstructive pulmonary disease Pulm. Pharmacol Ther. 2006;19:189–99.

22. Holford NHG, Sheiner LB. Kinetics of pharmacologic response. Pharmacol Ther. 1982;16:143–66.

23. Dyneka NL, Garg V, Jusko WJ. Comparison of four basic models of indirect pharmacologic responses. J Pharmacokinet Biopharmacodyn. 1993;21:457–78.

24. Kryzanski W, Jusko WJ. Mathematical formalism for the properties of four basic models of indirect pharmacodynamic responses. J Pharmacokinet Biopharmacodyn. 1997;25:107–23.

25. Exarchos TP, Sakellarios A, Siogkas PK, Fotiadis DI, Milosevic Z, Nikolic D, et al. Patient specific multiscale modelling for plaque formation and progression. Conf Proc IEEE Eng Med Biol Soc. 2012:2893–6.

26. Salvi S, Brashier B, Gothi D. Bronchodilator efficacy of tiotropium-formoterol via single pressurized meter dose inhaler (pmdi) versus tiotropium alone in COPD. Pulm Pharmacol Ther. 2013;13:130–2.

27. Kuna P, Ivanov Y, Trofimov VI, Saito T. Efficacy and safety of AZD3199 vs formoterol in COPD: a randomized, double-blind study. Respir Res. 2013;4(1):64.

28. Aryal S, Diaz-Guzman E, Mannino DM. COPD and gender differences: an update. Transl Res. 2013;162(4):208–18.

29. Cazzola M, Rogliani P, Ruggeri P, Segreti A, Proietto A, Picciolo S, et al. Chronic treatment with indacaterol and airway response to salbutamol in stable COPD. Respir Med. 2013;107(6):848–53.

30. Cazzola M, Page C, Matera MG. Long-acting muscarinic receptor antagonists for the treatment of respiratory disease. Pulm Pharmacol Ther. 2013;26:307–17.

31. Dolan S, Varkey B. Prognostic factors in chronic obstructive pulmonary disease. Curr Opin Pulm Med. 2005;11:149–52.

32. Han J, Dai L, Zhong N. Indacaterol on dyspnea in chronic obstructive pulmonary disease: a systematic review and meta-analysis of randomized placebo-controlled trials. BMC Pulm Med. 2013;25: 13–26.

33. Exposure-Response Working Group, Medical Policy Coordinating Committee. Guidance for industry. Exposure-response relationships — study design, data analysis, and regulatory applications. Food and Drug Administration, USA, 2003. Available from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072109.pdf. Accessed on 7 Jan 2014

34. Woodruff PG. Novel outcomes and end points: biomarkers in chronic obstructive pulmonary disease clinical trials. Proc Am Thorac Soc. 2011;8:350–5.