ARTICLE

Contribution of machine learning to tumor growth inhibition modeling for hepatocellular carcinoma patients under Roblitinib (FGF401) drug treatment

Mélanie Wilbaux[1] | **David Demanse**[2] | **Yi Gu**[3] | **Astrid Jullion**[2] | **Andrea Myers**[4] | **Vasiliki Katsanou**[5] | **Christophe Meille**[2] ⬤

[1]Pharmacometrics, Novartis, Basel, Switzerland

[2]Early Development Analytics, Novartis, Basel, Switzerland

[3]Pharmacokinetic Sciences, Novartis Institutes for Biomedical Research, Cambridge, USA

[4]Clinics, Novartis, Shanghai, China

[5]Clinical Operations, Novartis, Basel, Switzerland

Correspondence
Christophe Meille, Novartis Pharma AG Novartis Campus, Fabrikstrasse 2, CH-4056 Basel, Switzerland.
Email: christophe.meille@novartis.com

Abstract

Machine learning (ML) opens new perspectives in identifying predictive factors of efficacy among a large number of patients' characteristics in oncology studies. The objective of this work was to combine ML with population pharmacokinetic/ pharmacodynamic (PK/PD) modeling of tumor growth inhibition to understand the sources of variability between patients and therefore improve model predictions to support drug development decisions. Data from 127 patients with hepatocellular carcinoma enrolled in a phase I/II study evaluating once-daily oral doses of the fibroblast growth factor receptor FGFR4 kinase inhibitor, Roblitinib (FGF401), were used. Roblitinib PKs was best described by a two-compartment model with a delayed zero-order absorption and linear elimination. Clinical efficacy using the longitudinal sum of the longest lesion diameter data was described with a population PK/PD model of tumor growth inhibition including resistance to treatment. ML, applying elastic net modeling of time to progression data, was associated with cross-validation, and allowed to derive a composite predictive risk score from a set of 75 patients' baseline characteristics. The two approaches were combined by testing the inclusion of the continuous risk score as a covariate on PD model parameters. The score was found as a significant covariate on the resistance parameter and resulted in 19% reduction of its variability, and 32% variability reduction on the average dose for stasis. The final PK/PD model was used to simulate effect of patients' characteristics on tumor growth inhibition profiles. The proposed methodology can be used to support drug development decisions, especially when large interpatient variability is observed.

Mélanie Wilbaux and David Demanse contributed equally to this work.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Oncology trials, especially in phase I, are characterized by high patients' heterogeneity and variability in response, implying the need for a deep understanding of potential contributing factors.

WHAT QUESTION DID THIS STUDY ADDRESS?
Can machine learning be combined with population pharmacokinetic/pharmacodynamic (PK/PD) modeling to explain the high variability observed in clinical trials?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
A new approach combining machine learning and PK/PD modeling is proposed to improve model predictions by including a large set of patients' baseline characteristics.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?
This approach is susceptible to detect unexpected factors as covariates of interest in PK/PD models, supporting the definition of inclusion and exclusion criteria in future studies and could pave the way for precision medicine and individualized therapy.

INTRODUCTION

Oncology trials, especially in phase I, are associated with high patients' heterogeneity (e.g., inclusion of patients with multiple previous lines of treatments), small sample size, and rapid escalation in dose. Large variability in response is also usually observed implying the need for a deep understanding of potential contributing factors.

Pharmacometrics modeling and simulation (M&S) approaches are widely used in pharmaceutical industries to understand drug exposure-response relationships and to support drug development decisions, such as dose finding. Pharmacokinetic/pharmacodynamic (PK/PD) modeling of tumor growth inhibition (TGI) has been extensively applied to identify the optimal dose.[1] Nevertheless, patients' heterogeneity is reflected in the high interpatient variability of model parameters and can limit the usefulness of model simulations.

Over the past decade, progresses in technology and growing automation enabled companies to collect vast amounts of data. Machine learning (ML) is a subfield within artificial intelligence (AI), which aims to develop and utilize algorithms that learn from big data and then make a prediction about the future state of any new data set.[2,3] ML is gaining increasing interest within the pharmaceutical industry as it can promote data-driven decision making and has the potential to accelerate and reduce failure rates in drug discovery and development. Opportunities to apply such approaches occur in all stages of drug discovery and development, from identification of novel targets to analysis of digital pathology data in clinical trials.[4–6] ML methods are also being applied within the healthcare setting leading to significant advances in personalized medicine.[7]

Additionally, ML opens new perspectives to complement M&S approaches with a potential for combining both approaches to improve model performance, personalization, and predictivity, and make informed decisions from all available patients' data using robust algorithms.[8–14]

The main objective of this paper is to share a methodology combining ML with population PK/PD modeling to understand and characterize the sources of variability between patients and therefore improve model predictions for drug development support. For that, we used data from the phase I/II study evaluating the daily oral administration of Roblitinib (FGF401) in hepatocellular carcinoma (HCC) and other solid tumors.[15] Fibroblast growth factor (FGF)401 is a fibroblast growth factor receptor (FGFR)4 inhibitor, mainly indicated for the treatment of HCC, a deadly disease with limited treatment options, for which the FGF19–FGFR4–β Klotho (KLB) signaling pathway is a key driver.[16]

The relationship between FGF401 exposure and clinical efficacy was explored with a population PK/PD model of TGI, using tumor size expressed as the sum of the longest diameter (SLD). ML was used to derive a composite score of baseline factors predictive of time to progression (TTP), which was subsequently included in the PK/PD model.

METHODS

The proposed methodology along with the data used to combine ML with population PK/PD modeling of TGI is presented in Figure 1.

Data

Data from the first-in-human (FIH) phase I/II study evaluating oral administration of Roblitinib in HCC and other solid tumors (NCT02325739) were used.[15] The study protocol and all amendments were reviewed by the independent ethics committee or institutional review board at each study center. The study, including data collection, was conducted according to the International Conference on Harmonization (ICH) E6 Guidelines for Good Clinical Practice, which have their origin in the Declaration of Helsinki. This study was designed as a multicenter, open-label, nonrandomized study starting with a phase I dose escalation part followed by a phase II part. In the phase I part, the escalated Roblitinib dose regimens included daily (q.d.) administration of 50 mg, 80 mg, 120 mg, and 150 mg under fasted condition, and 80 mg and 120 mg under fed condition. Patients enrolled in the phase II part received Roblitinib at 120 mg q.d. under fasted condition, which was the recommended dose for expansion defined in the phase I part. A complete treatment cycle was defined as 21 days of continuous dosing.

In the phase I part, 10 PK blood samples per individual were collected on the first and the eighth day of cycle 1, and on the first day of cycle 2. In the phase II part, four PK blood samples were collected on the first day of cycle 1 and cycle 2. Several trough samples were also collected during the study.

Tumor size assessments, expressed as SLD, were scheduled at baseline and subsequently every 6 weeks (± 7 days) from cycle 3 until disease progression.

The TTP, used for the ML analysis, was defined as the time from the start of treatment date to the first documented progression or death due to underlying cancer. If a patient had no event, TTP was censored at the date of the last adequate tumor assessment.

SLD was chosen to account for the longitudinal profile of the drug effect on the tumor. TTP was selected as it was the primary end point from the trial and the closest PD end point related to tumor size.

PK data from all patients, whatever the tumor type, were first analyzed. Then, as HCC was the main indication for Roblitinib, patients with other tumor types were excluded for PK/PD modeling of TGI and ML analysis. Only data from patients with at least one PK sample and one post-treatment SLD measurement were kept in the analyses.

The limit of quantification (LOQ) was equal to 1.5 ng/ml for Roblitinib PK samples. The LOQ value for SLD data was assumed to be equal to the minimal value observed among patients. Concentrations and SLD values below the LOQ (BLOQ) were set to the LOQ values and flagged as left-censored observations in the dataset. BLOQ values were automatically taken into account in the likelihood as the probability of having an observation within an interval.

Two sets of baseline patients' characteristics were defined: (i) "basic set" including seven commonly tested covariates (i.e., height, weight, body mass index [BMI], age, race, sex, and food); and (ii) "extended set" including 75 individual factors (demographics, disease history, SLD, and laboratory data listed in Table 1). The basic set was used to evaluate covariates in the PK and TGI PK/PD models. The extensive set was used for the ML analysis.
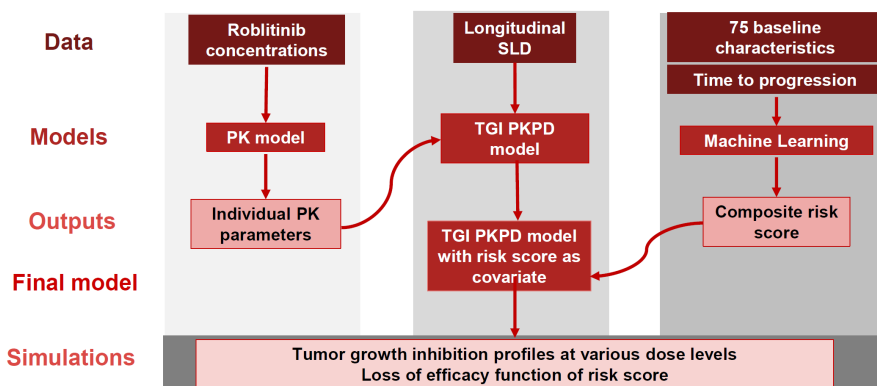

FIGURE 1 Contribution of ML to PK/PD tumor growth inhibition modeling for patients with hepatocellular carcinoma under Roblitinib drug treatment. A population PK model was first developed to describe Roblitinib concentrations. Then, the estimated individual PK parameters were used as inputs for the population PK/PD TGI model to described longitudinal SLDs data. ML was used to derive a composite score of baseline factors predictive of time to progression, which was subsequently included in the PK/PD model. FGF, fibroblast growth factor; ML, machine learning; PD, pharmacodynamic; PK, pharmacokinetic; SLD, sum of the longest diameters; TGI, tumor growth inhibition.

TABLE 1 Descriptive summary of PK and SLD data and patients' characteristics

Characteristics	Median [minimum − maximum]/number of individuals or observations [%]	
Data for PK model development		
Number of individuals	160	
Number of PK observations	3036	
Number of PK observations per individual	16 [1–39]	
Number of BLOQ values [LOQ = 1.5 ngml]	175 [6%]	
Number of individuals and PK observations per first dose levels:	Number of individuals	Number of observations
50 mg	11	217
80 mg	11	317
120 mg	131	2304
150 mg	7	198
Number of doses per individual	65 [1–808]	
Time of follow up of PK concentrations, days	61 [0.08–650.92]	
Basic covariates		
Height, cm	170 [149–186]	
Weight, kg	68.5 [38.0–114.0]	
BMI, kg/m^2	24.2 [14.2–35.2]	
Age, years	62 [21–85]	
Number of individuals per race:		
Non-Asian	87 [54%]	
Asian	73 [46%]	
Number of individuals per gender		
Male	118 [74%]	
Female	42 [26%]	
Number of individuals with food		
Fasted	136 [85%]	
Fed	24 [15%]	
Data for PK/PD tumor growth inhibition model development		
Number of individuals	127	
Number of SLD observations	511	
Number of SLD observations per individual	3 [1–20]	
Number of BLOQ values [LOQ = 9 mm]	6 [1.2%]	
Number of individuals and SLD observations per first dose levels	Number of individuals	Number of observations
50 mg	8	22
80 mg	10	43
120 mg	103	406
150 mg	6	40
Time of follow-up, months	2.7 [0–26.3]	
Baseline SLD, mm	92 [15–352]	

Abbreviations: BLOQ, below the limit of quantification; BMI, body mass index; LOQ, limit of quantification; PD, pharmacodynamic; PK, pharmacokinetic; SLD, sum of the longest diameters.

Model development

Tumor growth inhibition PK/PD model

A minimal mathematical PK/PD model was developed to establish the relationship between Roblitinib exposure and longitudinal profiles of SLD. Population typical parameters and their interindividual variability (IIV) were estimated based on a population analysis applying a non-linear mixed-effect modeling approach.[17]

A sequential approach was applied, with the estimated individual PK parameters (mode of the conditional parameter distribution for each individual) from the final population PK model used as inputs for the population PK/PD TGI model.

Several PK structural models were tested including one, two, or three compartments, with first- or zero-order absorption, with or without absorption delay.[18]

For the tumor growth inhibition, several structural models were investigated to describe longitudinal SLD data starting from a standard TGI structural model that combines tumor growth and killing rates.[19,20] Various models were tested for unperturbed tumor growth, such as exponential and linear growth. To reproduce the observed delay in drug effect, effect and transit compartments were evaluated. Drug effect on tumor growth was investigated with various parametrizations and inclusion of a resistance term was tested to reproduce tumor regrowth under treatment.[21] The model was parametrized to allow direct estimation of an individual dose for tumor stasis instead of drug potency.

In both PK and PK/PD models, visual inspection of the potential covariate relationships was investigated by plotting the conditional distribution of each parameter versus covariates.[22] Selection of covariate-parameter relationship to be retained in the final model was based on several criteria: difference in Akaike information criterion (AIC), reduction of IIV, and significant p values ($p \leq 0.05$) for the Wald test on model parameters. The clinical relevance of the statistically significant PK covariates was evaluated by simulating their effect on PK exposure (e.g., area under the curve [AUC] and maximum plasma concentration [C_{max}]).

PK and PK/PD model evaluation and selection

Model evaluation and selection were assessed using the following criteria: Stochastic Approximation Expectation–Maximization (SAEM) algorithm convergence assessment, change in the objective function value and AIC, precision of the parameter estimates (relative standard error [RSE]) and degree of correlation between them, decreases in both IIV and residual variability, and visual inspection of diagnostic plots. Diagnostic plots were assessed to evaluate model adequacy and possible lack of fit or violation of assumptions. The predictive performance of the final model was evaluated by simulating data using parameter estimates (fixed and random effects) and plotting visual predictive check (VPC). Prediction-corrected VPCs (pcVPCs) were provided, to remove the baseline variability coming from the various doses and/or potential covariates.[23]

Application of machine learning to identify baseline predictive factors of progression

A survival Cox proportional hazards model was applied on the individual TTP end point. The following equation was used to describe the instantaneous risk of an event at time t for an individual i with a set of p baseline characteristics $x_i = (x_{i,1}, \ldots , x_{i,p})$:

$$h(t|x_i) = h_0(t) \cdot e^{x_i \cdot \beta}$$

where $h_0(t)$ is the baseline hazard representing the hazard when all o te predictors $\beta = (\beta_1, \ldots , \beta_p)$ are equal to zero. The predicted hazard $h(t|x_i)$ is the product of the baseline hazard and the exponential function of the linear combination of the predictors.

The extended set of 75 baseline individual factors was used to evaluate predictors of TTP. Each variable was standardized to have a mean of zero and an SD of one.

First, a univariate Cox regression analysis was performed to evaluate the effect of each baseline characteristic on survival. Data were split using random sampling into a training set of 80% and test set of 20% of the total number of patients. Then, a supervised ML approach, using a penalized Cox regression, in combination with cross-validation (leave-one-out cross-validation) methodology was implemented using training set data. This approach allows to obtain reliable survival prediction models for high-dimensional predictors.[24] The penalized likelihood approach used to take into account correlated predictors and solve multicollinearity was the elastic net that integrates the ridge and the lasso penalties.[25] Root mean squared error (RMSE) criteria was used to select the optimal tuning penalties (alpha and lambda) for the elastic net model during the cross-validation part. Then, model prediction performance for the final elastic net model was assessed using the time dependent receiver operating characteristic (ROC) curves on the test set.[26]

The elastic net model coefficients using the best penalties during the tuning on the training part were used to derive a composite risk function and applied on each patient (training and test sets) to obtain a continuous risk score value. To facilitate clinical interpretation, the risk score was used to categorize the population into two subgroups based on the median

value: (i) patients at low risk of progression if the score was below the median, and (ii) patients at high risk of progression if the score was above or equal to the median.

Combination of machine learning with PK/PD modeling

The selected approach to combine ML with PK/PD modeling was to evaluate the continuous risk score as a potential covariate of interest on PD parameters. Redundancy between baseline covariates and factors from the risk score was avoided by excluding them from the multivariate analysis. Selection of parameter-score relationship to be retained in the final model was based on the same criteria as the classical approach: AIC improvement, IIV reduction, significant p values ($p \leq 0.05$) for the Wald test on model parameters.

Model simulations

The final PK/PD model including the risk score derived from ML was used to simulate FG401-induced tumor growth inhibition at the observed dose levels given patients' baseline characteristics (500 simulations of the available dataset). Quantitative comparisons of the various dose levels were assessed by simulations of the proportion of responders by risk group at 4 and 6 months (based on clinical considerations) after the start of treatment. Simulated patients were considered as responders when SLD change from baseline was lower than +20% at the time considered.[27]

The final PK/PD model was also used to simulate, across 5000 virtual patients, the time to 50% loss of drug potency function of the risk score in the range of patients' study values.

Computing process

Population PK model development was performed using the MonolixSuite2018.R1 software and population PK/PD model using the MonolixSuite2019.R2 version.[28,29] R version 3.4.3 software was used for data and outputs processing and R version 3.6.1 for simulations.[30] Simulations were performed using *simulx* function from the R package

mlxR_4.2.0. Elastic net model was performed in R version 3.4.3 software using packages caret version 6–0.84 and *glmnet* version 2.0–18.

RESULTS

Data

The analysis of Roblitinib PK included 3036 observations from 160 individuals with HCC and other solid tumors. A total of 50 PK samples has been initially excluded due to: (i) abnormal PK concentration (e.g., abnormally high trough concentrations), (ii) predose sample collected postdosing, or (iii) vomiting within 4 h of drug administration.

A total of 33 individuals among the 160 (21%) were excluded due to non-HCC tumor type for the PK/PD and ML analyses. The longitudinal analysis of SLD data used 511 observations from 127 patients with HCC. Each patient had at least one PK sample and one SLD measurement. A descriptive summary of these data and the basic covariates are given in Table 1.

Individual SLD kinetics profiles by dose group are shown in Figure 2. Large variability between individual profiles was observed.

The minimal SLD value observed among patients was equal to 9 mm and was assumed to correspond to the LOQ.

A total of 103 progression events in the 127 patients with HCC (81%) were observed, with a median TTP of 2.92 months (95% confidence interval [CI]: 2.66 months – 4.07 months).

Selected PK/PD model of tumor growth inhibition

A two-compartment model with a delayed zero-order absorption and linear elimination was adequately describing Roblitinib PK data. The structural PK model is summarized in Figure 3 and is described by the following equations for one given patient, with k corresponding to the dose administration number and the time of first administration being $t_{D,1} = 0$:

$$t_{D,k} < t \leq t_{D,k} + \text{Tlag}: \frac{dA_1}{dt} = 0 \qquad\qquad A_1(0) = 0$$

$$t_{D,k} + \text{Tlag} < t \leq t_{D,k} + \text{Tlag} + Tk0: \frac{dA_1}{dt} = \frac{\text{Dose}_k}{Tk0} - \frac{Cl}{V_1} \cdot A_1(t) + k_{21} \cdot A_2(t) - k_{12} \cdot A_1(t) \quad A_1(0) = 0$$

$$\text{Tlag} + Tk0 + t_{D,k} < t \leq \text{Tlag} + Tk0 + t_{D,k+1}: \frac{dA_1}{dt} = -\frac{Cl}{V_1} \cdot A_1(t) + k_{21} \cdot A_2(t) - k_{12} \cdot A_1(t) \quad A_1(0) = 0$$

$$\forall t: \frac{dA_2}{dt} = k_{12} \cdot A_1(t) - k_{21} \cdot A_2(t) \qquad\qquad A_2(0) = 0$$
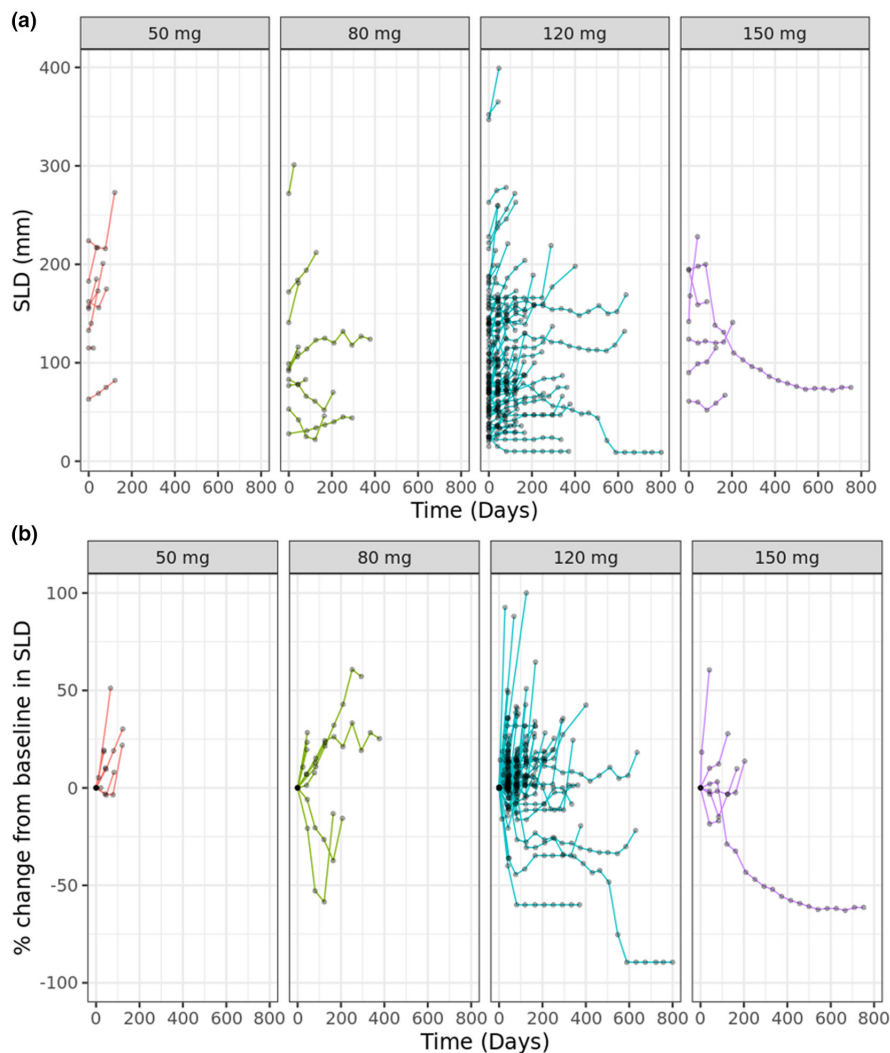
(a)

(b)

FIGURE 2 Observed individual SLD kinetics profiles by dose levels. Individual SLD kinetics profiles (a) raw data and (b) percent change from baseline stratified by dose group. SLD, sum of the longest diameters.

Considering:

$$Q = k_{12} \cdot V_1$$
$$V_2 = \frac{k_{12}}{k_{21}} \cdot V_1$$

With:

- $A_1(t)$, the Roblitinib amount in the central compartment at time t
- $A_2(t)$, the Roblitinib amount in the peripheral compartment at time t
- $t_{D,k}$, the time of kth dose administration
- Tlag, the delay before the absorption
- $Tk0$, the duration of the zero-order absorption
- Cl, the clearance
- V_1, the volume of distribution
- Q, the intercompartmental clearance
- V_2, the volume of distribution of the peripheral compartment.

The apparent clearance (Cl/F) was estimated at $19.7\,\mathrm{L\,h^{-1}}$ and apparent volume of distribution (V_1/F) at $110\,\mathrm{L}$. Six covariates from the basic set were found significant on key PK model parameters. Food, high Roblitinib dose, and low BMI were associated with a longer absorption. Gender and race (Asian vs. non-Asian) were found as statistically significant covariates on Cl/F and V_1/F, with lower values predicted in females and Asians. Increasing baseline weight was associated with increasing V_1/F. None of the covariates had a clinically meaningful influence on Roblitinib exposure metrics. Parameter estimates from the final PK model are reported in Table 2.

Unperturbed tumor growth was best characterized by an exponential function with a growth parameter kg. Predicted Roblitinib individual plasma PK profiles were linked to the tumor-killing rate parametrization through an effect compartment to reproduce the observed delay before drug effect. A resistance component, with parameter λ, was added to describe the tumor regrowth under treatment.[21] The structural TGI model, is summarized in Figure 3 and is described by the following ODE equations:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathrm{ke}_0 \cdot C(t) - \mathrm{ke}_0 \cdot E(t) \qquad E(0) = 0$$
$$\frac{\mathrm{dSLD}}{\mathrm{d}t} = \mathrm{kg} \cdot \mathrm{SLD}(t) - \mathrm{kr}(t) \cdot \mathrm{SLD}(t) \cdot e^{-\lambda \cdot t} \quad \mathrm{SLD}(0) = \mathrm{SLD}_0$$
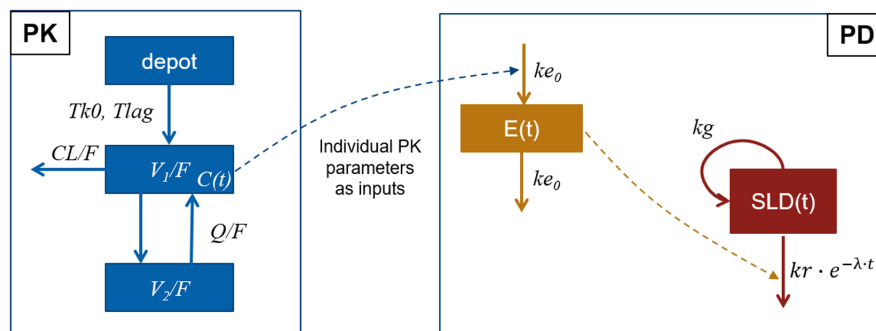
FIGURE 3 Structural PK/PD tumor growth inhibition model. Tlag (h^{-1}) is the delay before the absorption. Tk0 (h) is the duration of the 0-order absorption. $C(t)$ (ng/ml) corresponds to the drug concentration in the central compartment at time t (h). Cl/F ($L\,h^{-1}$) is the apparent clearance. V_1/F (L) is the apparent volume of distribution. Q/F ($L\,h^{-1}$) is the apparent intercompartmental clearance. V_2/F (L) is the apparent volume of distribution of the peripheral compartment. $E(t)$ represents the concentration in the effect compartment at time t. ke_0 (h^{-1}) is the transit rate of the effect compartment. $SLD(t)$ (mm) represents the sum of the longest diameters at time t. kg (h^{-1}) is the tumor growth rate. kr (h^{-1}) is the drug killing rate. λ (h^{-1}) is the resistance parameter. PD, pharmacodynamic; PK, pharmacokinetic.

With:

- $E(t)$, the concentration in the effect compartment at time t
- ke_0, the transit rate
- $C(t)$, the drug concentration at time t
- kg, the tumor growth rate
- $SLD(t)$, the SLD at time t
- $kr(t)$, the drug killing rate at time t
- λ, the resistance parameter
- SLD_0, the baseline SLD.

The killing rate at time t, assuming a linear PK and for an administration period τ, was calculated as:

$$kr(t) = \frac{E(t) \cdot Cl/F \cdot kg \cdot \tau}{\text{avDS}}$$

With:

- Cl/F, the apparent Roblitinib clearance
- τ, the dosing interval (24 h)
- avDS, the average dose for tumor stasis in absence of resistance.

None of the basic covariates was found significant on any PD model parameters.

Univariate Cox regression and Elastic net models

The resulted forest plot from the univariate Cox regression analysis using the 75 patients' baseline characteristics is available in Figure S1.

Elastic net model with cross-validation methodology resulted in four baseline predictive factors of the risk of progression: age, number of metastases, lymphocyte count, and portal vein invasion (PVI). Of note, individual PK parameters from the final population PK model were also tested. The individual apparent volume of distribution $V_{1,i}/F_i$ was found as a predictive factor, indicating a link between the PK and TTP (results shown in Material S1). As individual PK parameters were already part of the PK/PD model, $V_{1,i}/F_i$ was removed from the risk function. Parameter estimates from the univariate Cox models are shown in Table 3. The predictive accuracy of the elastic net model is available in Table S3. The accuracy was 67% in the training set and 64% in the testing set at 2 months.

The following composite risk function was derived to calculate the continuous risk score for each patient:

$$\text{Score}_i = 0.197 - 0.184 \cdot \text{lymphocytes}_i{}^* - 0.044 \cdot \text{age}_i{}^* + 0.027 \cdot \text{metastases}_i{}^* + 0.024 \cdot \text{PVI}_i{}^*$$

Each individual predictive factor, factor*, was standardized to have a mean of zero and a SD of one. Mean and SD for each factor from the training set are reported in Table S2.

The lymphocyte count was found to have the highest effect on TTP, followed by age, number of metastases and PVI. A median score of 0.22 was observed in the population with first quartile at 0.06 and third quartile at 0.36 (Figure S2). Patients were stratified into two groups: (i) patients at high risk of progression if score$_i$ was greater than or equal to 0.22, and (ii) patients at low risk of progression if score$_i$ was less than or equal to 0.22. The survival curves stratified by risk group are available in Figure S3.

Parameter (unit)	Fixed effect (% RSE)	IIV: SD of random effect (% RSE)	p value Wald test
PK model			
Tlag (h^{-1})	0.268 (11)	0.63 (12)	/
Tk0 (h) (fasted condition)	0.811 (10)	0.64 (10)	/
Tk0 in fed condition (h)	1.58 (25)	/	5×10^{-5}
BMI effect on Tk0	−1.66 (25)	/	7×10^{-5}
Dose effect on Tk0	0.983 (28)	/	4×10^{-4}
Cl/F ($L\,h^{-1}$) (non-Asian male)	19.7 (4)	0.29 (6)	/
Cl/F in Asian	16.3 (26)	/	1×10^{-4}
Cl/F in female	15.1 (21)	/	2×10^{-6}
V_1/F (L) (non-Asian male)	110 (3)	0.15 (12)	/
Weight effect on V_1/F	0.332 (33)	/	0.002
V_1/F in female	84.5 (20)	/	5×10^{-7}
V_1/F in Asian	84.5 (17)	/	2×10^{-9}
Q/F ($L\,h^{-1}$)	5.59 (6)	0 FIX	/
V_2/F (L)	49.2 (7)	0.68 (9)	/
Combined residual error: additive (ng/ml)/proportional (%)	4.1 (7)/33 (2)	/	/
Tumor growth inhibition model			
ke_0 (h^{-1})	7.79×10^{-4} (33)	1.50 FIX	/
SLD0 (mm)	89.95 (6)	0.62 (6)	/
λ (h^{-1})	3.40×10^{-4} (30)	1.15 (16)	/
β_{Score}	2.87 (32)	/	0.002
kg (h^{-1})	1.05×10^{-4} (12)	0.78 (14)	/
avDS (mg)	59.42 (24)	0.53 (28)	/
Residual additive error (mm)	6.44 (5)	/	/

TABLE 2 Parameter estimates of the final PK/PD tumor growth inhibition model using a sequential PK/PD modeling approach

Abbreviations: BMI, body mass index; FIX, fixed parameter; IIV, interindividual variability; PD, pharmacodynamic; PK, pharmacokinetic; RSE, relative standard error.

Final tumor growth inhibition model

The continuous risk score was evaluated as a covariate on PD parameters from the TGI model and found significant on the resistance parameter, λ, with an increasing score associated with an increasing typical λ value. The score effect on the resistance parameter was implemented using the following equations:

$$\lambda_i = \lambda_{pop} \cdot e^{\beta_{Score} \cdot (Score_i - 0.22)}$$

With:

- λ_i, the resistance parameter value for the individual i
- λ_{pop}, the typical resistance parameter value in the population
- β_{Score}, the score effect on the resistance parameter
- $Score_i$, the risk score value for the individual i.

TABLE 3 Parameter estimates of the univariate models

Predictive factor	Summary (mean [SD]/number of individuals [%])	HR (univariate)
Age, years	61.5 [12.1]	0.98 (0.97–1.00, $p = 0.016$)
Number of metastases	1.9 [1.0]	1.23 (1.02–1.49, $p = 0.033$)
Lymphocyte count (10^9 cells L^{-1})	1.1 [0.5]	0.43 (0.29–0.65, $p < 0.01$)
Portal vein invasion		
Yes	27 [21.4]	2.22 (1.35–3.66, $p = 0.002$)
No	99 [78.6]	
$V_{1,i}/F_i$ (L)	93.4 [22.0]	0.99 (0.98–1.00, $p = 0.006$)

Abbreviations: HR, hazard ratio; p, p value; $V_{1,i}/F_i$, volume of distribution from the population PK model.

TABLE 4 Simulated proportions of responders at 4 and 6 months[a] in patients at high and low risk of progression (median [25th–75th quantiles])

End point/dose	50 mg	80 mg	120 mg	150 mg
High risk				
% Patients at 4 months	40.6 [37.5–45.3]	50.0 [46.9–54.7]	59.4 [54.7–64.1]	64.1 [59.4–67.6]
% Patients at 6 months	25.0 [20.3–28.1]	34.4 [29.7–37.5]	43.8 [39.1–48.4]	50.0 [43.8–53.1]
Low risk				
% Patients at 4 months	49.2 [44.4–54.0]	60.3 [57.1–65.1]	69.8 [66.7–74.6]	74.6 [71.4–79.4]
% Patients at 6 months	33.3 [30.2–38.1]	47.6 [42.9–52.4]	60.3 [55.6–63.5]	66.7 [61.9–69.8]

Note: Patients with SLD change from baseline lower than +20% (including stable disease) were considered as responders.

Abbreviation: SLD, sum of the longest diameters.

[a]Assuming months of 30 days.

Inclusion of the risk score as a covariate on λ_{pop} resulted in reduction of the SD of its random effect by 19% (decrease from 1.42 to 1.15). The SD of the random effect on the average dose for tumor stasis (avDS) was also reduced by 32% (decrease from 0.79 to 0.53). Parameter estimates from the final population PK/PD model of tumor growth inhibition are reported in Table 2. Uncertainty on estimated parameters were considered as acceptable, with all RSE less than or equal to 32%. IIV on all PD parameters were relatively high, reflecting the observed heterogeneity in spaghetti plots. Variability on the parameter ke_0, the transit rate from the effect compartment that reproduces the delay in drug effect, was fixed to a high value (>100%) to obtain stable model parameter estimation with SAEM convergence achieved for all parameters. Simpler structural models, with no delay component, were resulting in worst AIC and goodness-of-fit plots.

Model evaluation

According to goodness-of-fit plots, VPCs and pcVPC, Roblitinib PK and change in SLD data were adequately fitted by the final models with no apparent bias (Figures S4–S10).

Model simulations

The proportion of responders at 4 and 6 months in both high-risk and low-risk patients are provided in Table 4. These simulations reproduced a dose-dependent tumor growth inhibition. A dose of 120 mg q.d. leads to a response (median [25th–75th quantiles]) at 4 and 6 months in 59.4% [54.7%–64.1%] and 43.8% [39.1%–48.4%] of patients at high risk, and in 69.8% [66.7%–74.6%] and 60.3% [55.6%–63.5%] of patients at low risk, respectively.

Interestingly, a typical patient with a median risk score is predicted to show 50% loss of drug efficacy at (median [10th–90th percentile]) 81.5 days [19.7 days–352.0 days] (Figure S11).

DISCUSSION

The present analysis illustrates the combination of ML with a population PK/PD model to improve model prediction. First, a population PK model was developed to derive individual PK parameters that were used as inputs for the population PK/PD model of tumor growth inhibition in patients with HCC. Then, a penalized Cox regression model was applied to evaluate baseline predictors of TTP and derive a composite risk score used to supplement the population PK/PD model. For the purpose of that analysis, the elastic net regression was selected among other linear and nonlinear ML algorithms (e.g., random forest, neural net, and support vector machine) based on interpretability and parsimony. Four baseline factors predictive of TTP were identified among the 75 patients' characteristics. Decreasing age and lymphocyte count, increasing number of metastases, as well as PVI were associated with an increasing risk of progression. Even if the number of metastases and PVI were expected to influence TTP,[31,32] the lymphocyte count had the highest effect. Evidence suggests that lymphocytes play a key role in the development and progression in HCC, as liver cancer arises upon chronic hepatic inflammation.[33] However, inflammation effect is still controversial as it exerts both pro- and anti-tumor effects, and the associated molecular and cellular cascades are still not well understood.[33] One hypothesis for the observed decreasing age associated with increasing risk of progression would be a lower tumor growth in elderly patients, a trend found with the parameter *kg* in the PK/PD model without score, but that was not statistically significant. Another hypothesis would be a potential

selection bias, as patients are part of a trial population that could result in younger patients with worse disease and older patients with less aggressive disease.[34]

Genetic information data, such as patients' mutation profile, was not available for the current analysis, but genetic influence on TTP could be considered to inform future studies and optimize recruitment criteria on top of the four baseline predictive factors.

Results from the penalized Cox regression model were combined with the TGI model by including the derived continuous risk score as a covariate on the PD model resistance parameter λ. The IIV on λ and on the average dose for stasis, $avDS$, was reduced by 19% and 32%, respectively, and the score was statistically significant (p value from the Wald test $p = 0.002$). Parameters of the model were well-estimated, but large unexplained variability was found on all PD parameters, including λ. Prediction intervals of the pcVPC are wide and shows that observations were not well captured beyond 8 months; the integration of a dropout model could be investigated to potentially solve this bias. Influence of the risk score on the resistance parameter λ could be expected as this parameter reflects the exponential loss of efficacy in function of time and the resulting regrowth of the tumor.

The score derivation is dependent on the observed TTP, which is a post-baseline end point. Consequently, the use of the score as a covariate on the λ parameter is derived from baseline patients' characteristics but influenced by a post-baseline observation. However, even if the ML end point (TTP) and PK/PD model end point (SLD) were not similar, the risk score appeared to be a significant covariate in the PK/PD model.

The predictive accuracy of the elastic net model is associated with a poor discrimination, that can be explained by the absence of dose and PK information in the composite risk score. The PK information was excluded to avoid inclusion of a post baseline parameter and redundancy with the PK/PD model. However, this composite risk score appears to be useful and relevant, as it was found to be highly significant on the PK/PD resistance parameter, and to explain a part of its variability.

The final PK/PD model including the risk score was used to simulate the dose-dependent TGI profiles. Nevertheless, the developed PK/PD model does not reflect the full Response Evaluation Criteria in Solid Tumors (RECIST) criteria, and thus simulations of responders are only based on the simulated longitudinal SLD profiles and do not consider other aspects, such as apparition of new lesions.

It was not possible to validate the results on an independent dataset to assess the robustness of the findings. In addition, this modeling framework has been performed based on limited data from a homogenous disease population and some results, such as influence of the score on the resistance parameter, may not be applicable to other tumor types or drugs.

In this work, the composite risk score from the penalized Cox regression model is used to keep the potential benefit of the collinearity between factors. Integration of the factors from the risk score as separate covariates on the resistance parameter in the PK/PD model have been explored. Results showed that only the age effect was statistically significant (p value from Wald test $= 0.01$) and identifiable. In addition, the variability on the resistance parameter was higher compared to the model including the score (SD of random effect: 1.45 vs. 1.15). Nevertheless, evaluation of the relevance of these factors as independent covariates could have been an alternative approach to investigate. In this case, ML would then be used as a screening tool to reduce data dimension and identify covariates of interest to be explored in a classical covariate search.[35]

PK/PD TGI modeling has been already combined with ML by Zwep et al.[36] They applied the least absolute shrinkage and selection operator (LASSO) method on the empirical Bayes estimates-derived TGI parameters to identify high-dimensional genomic predictors. They also used a group LASSO regression model to identify pathways associated with treatment response or resistance. Their results assumed that there is no misspecification in the structural models, whereas in the present work, ML was applied on a clinical end point (TTP), which is independent on the structural PK/PD model. Other approaches may be considered to combine ML and PK/PD modeling. Sibieude et al. showed that ML methods for covariate search achieved a better predictive ability of empirical Bayes clearance compared to classical methods (SCM and COSSAC).[37] Other publications applied ML approach to predict PK parameters, for example, the nivolumab clearance based on baseline cytokine data.[38] However, in our proposed work, we combined ML with a PK/PD model of efficacy using a more challenging end point (SLD) that is associated with high variability. We could also imagine using ML approach, such as long short term memory (LSTM) to predict longitudinal outcome, or as published by Lu et al. to apply a neural-PK/PD model with which they predicted drug concentration and platelet dynamics.[39,40]

ML provides several opportunities to improve robustness of data-driven decision making across drug discovery and development in oncology.[41] Combining ML and PK/PD modeling is proposed to improve model predictions by including a large set of patients' baseline characteristics. The proposed methodology is not limited to efficacy but could also be applied for safety events. This approach is susceptible to detect unexpected factors as covariates of interest in PK/PD models, supporting the definition of inclusion and exclusion criteria in future studies and could

pave the way for precision medicine and individualized therapy.

CONFLICT OF INTEREST
All authors were employees of Novartis at the time of the manuscript submission.

ORCID
Christophe Meille https://orcid.org/0000-0001-8573-1248

REFERENCES

1. Park K. A review of modeling approaches to predict drug response in clinical oncology. *Yonsei Med J*. 2017;58:1-8.

2. Murphy KP. *Machine Learning: A Probabilistic Perspective*. MIT Press; 2012.

3. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.

4. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. *Drug Discov Today*. 2018;23:1241-1250.

5. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov*. 2019;18:463-477.

6. Talevi A, Morales JF, Hather G, et al. Machine learning in drug discovery and development part 1: a primer. *CPT Pharmacometrics Syst Pharmacol*. 2020;9:129-142.

7. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380:1347-1358.

8. Hutchinson L, Steiert B, Soubret A, et al. Models and machines: how deep learning will take clinical pharmacology to the next level. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:131-134.

9. Nicolò C, Périer C, Prague M, et al. Machine learning and mechanistic modeling for prediction of metastatic relapse in early-stage breast cancer. *JCO Clinical Cancer Informatics*. 2020;4:259-274.

10. Koch G, Pfister M, Daunhawer I, Wilbaux M, Wellmann S, Vogt JE. Pharmacometrics and machine learning partner to advance clinical data analysis. *Clin Pharmacol Ther*. 2020;107:926-933.

11. McComb M, Bies R, Ramanathan M. Machine learning in pharmacometrics: opportunities and challenges. *Br J Clin Pharmacol*. 2021;88:1482-1499. doi:10.1111/BCP.14801

12. Chan P, Zhou X, Wang N, Liu Q, Bruno R, Jin JY. Application of machine learning for tumor growth inhibition – overall survival modeling platform. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:59-66.

13. González-García I, Pierre V, Dubois VFS, et al. Early predictions of response and survival from a tumor dynamics model in patients with recurrent, metastatic head and neck squamous cell carcinoma treated with immunotherapy. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:230-240.

14. Benzekry S. Artificial intelligence and mechanistic modeling for clinical decision making in oncology. *Clin Pharmacol Therap*. 2020;108:471-486.

15. Chan SL, Yen CJ, Schuler M. et al. Abstract CT106: Ph I/II study of FGF401 in adult pts with HCC or solid tumors characterized by FGFR4/KLB expression. CT106–CT106; 2017. doi:10.1158/1538-7445.AM2017-CT106

16. Weiss A, Adler F, Buhles A, et al. FGF401, a first-in-class highly selective and potent FGFR4 inhibitor for the treatment of FGF19-driven hepatocellular cancer. *Mol Cancer Ther*. 2019;18:2194-2206.

17. Sheiner LB. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. *Drug Metab Rev*. 1984;15:153-171.

18. Mould DR, Upton RN. Basic concepts in population modeling, simulation, and model-based drug development—part 2: Introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e38.

19. Ribba B, Holford NH, Magni P, et al. A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT: Pharm Syst Pharmacol*. 2014;3(5):e113. doi:10.1038/psp.2014.12

20. Yin A, Moes DJAR, Hasselt JGC, Swen JJ, Guchelaar HJ. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:720-737.

21. Claret L, Girard P, Hoff PM, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. *J Clin Oncol*. 2009;27:4103-4108.

22. Lavielle M, Ribba B. Enhanced method for diagnosing Pharmacometric models: random sampling from conditional distributions. *Pharm Res*. 2016;33:2979-2988.

23. Bergstrand M, Hooker AC, Wallin JE, Karlsson MO. Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. *AAPS J*. 2011;13:143-151.

24. Bøvelstad HM, Nygård S, Størvold HL, et al. Predicting survival from microarray data a comparative study. *Bioinformatics*. 2007;23:2080-2087.

25. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol*. 2005;67:301-320.

26. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56:337-344.

27. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247.

28. Monolix version 2018R1. Lixoft SAS; 2018. http://lixoft.com/products/monolix/

29. Monolix version 2019R2. Lixoft SAS; 2019. http://lixoft.com/products/monolix/

30. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2017. https://www.R-project.org/

31. Chan SL, Chong CCN, Chan AWH, Poon DMC, Chok KSH. Management of hepatocellular carcinoma with portal vein tumor thrombosis: review and update at 2016. *World J Gastroenterol*. 2016;22:7289-7300.

32. Connolly GC, Chen R, Hyrien O, et al. Incidence, risk factors and consequences of portal vein and systemic thromboses in hepatocellular carcinoma. *Thromb Res*. 2008;122:299-306.

33. Mossanen JC, Tacke F. Role of lymphocytes in liver cancer. *Onco Targets Ther*. 2013;2:e26468.

34. Chang PE, Ong WC, Lui HF, Tan CK. Is the prognosis of young patients with hepatocellular carcinoma poorer than the prognosis of older patients? A comparative analysis of clinical characteristics, prognostic features, and survival outcome. *J Gastroenterol*. 2008;43:881-888.

35. Hutmacher MM, Kowalski KG. Covariate selection in pharmacometric analyses: a review of methods. *Br J Clin Pharmacol*. 2015;79:132-147.

36. Zwep LB, Duisters KLW, Jansen M, et al. Identification of high-dimensional omics-derived predictors for tumor growth dynamics using machine learning and pharmacometric modeling. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:350-361.

37. Sibieude E, Khandelwal A, Hesthaven JS, Girard P, Terranova N. Fast screening of covariates in population models empowered by machine learning. *J Pharmacokinet Pharmacodyn*. 2021;48:597-609.

38. Wang R, Zheng J, Shao X, et al. Development of a prognostic composite cytokine signature based on the correlation with nivolumab clearance: translational PK/PD analysis in patients with renal cell carcinoma. *J Immunother Cancer*. 2019;7:348.

39. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9:1735-1780.

40. Lu J, Bender B, Jin JY, Guan Y. Deep learning prediction of patient response time course from early data via neural-pharmacokinetic/pharmacodynamic modeling. *Nat Mach Intellig*. 2020;3:696-704.

41. Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. *Cancer Cell*. 2021;39:916-927.

SUPPORTING INFORMATION

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