Model-Based Biomarker Selection for Dose Individualization of Tyrosine-Kinase Inhibitors

Maddalena Centanni and Lena E. Friberg*

Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

Tyrosine-kinase inhibitors (TKIs) demonstrate high inter-individual variability with respect to safety and efficacy and would therefore benefit from dose or schedule adjustments. This study investigated the efficacy, safety, and economical aspects of alternative dosing options for sunitinib in gastro-intestinal stromal tumors (GIST) and axitinib in metastatic renal cell carcinoma (mRCC). Dose individualization based on drug concentration, adverse effects, and sVEGFR-3 was explored using a modeling framework connecting pharmacokinetic and pharmacodynamic models, as well as overall survival. Model-based simulations were performed to investigate four different scenarios: (I) the predicted value of high-dose pulsatile schedules to improve clinical outcomes as compared to regular daily dosing, (II) the potential of biomarkers for dose individualizations, such as drug concentrations, toxicity measurements, and the biomarker sVEGFR-3, (III) the cost-effectiveness of biomarker-guided dose-individualizations, and (IV) model-based dosing approaches versus standard sample-based methods to guide dose adjustments in clinical practice. Simulations from the axitinib and sunitinib frameworks suggest that weekly or once every two weeks high-dosing result in lower overall survival in patients with mRCC and GIST, compared to continuous daily dosing. Moreover, sVEGFR-3 appears a safe and cost-effective biomarker to guide dose adjustments and improve overall survival (€36 784.- per QALY). Model-based estimations were for biomarkers in general found to correctly predict dose adjustments similar to or more accurately than single clinical measurements and might therefore guide dose adjustments. A simulation framework represents a rapid and resource saving method to explore various propositions for dose and schedule adjustments of TKIs, while accounting for complicating factors such as circulating biomarker dynamics and inter-or intra-individual variability.

Keywords: tyrosine kinase inhibitor, dose individualization, population modeling, PKPD, pharmacometrics

INTRODUCTION

Tyrosine-kinase inhibitors (TKIs) represent a group of targeted therapies that have shown great benefit in the treatment of various malignancies (Arora and Scholar, 2005). Contrary to most cytotoxic drugs, TKIs can be taken orally, and they target fundamental proteins involved in cancer cell proliferation, differentiation, migration, and metabolism. Despite their advantages, TKIs demonstrate
high inter-individual variability in pharmacokinetics (PK) as well as in pharmacodynamics (PD), i.e. with respect to safety and efficacy outcomes, and, therefore, would benefit from dose individualization to optimize treatment outcomes (Sabanathan et al., 2017; Schmидinger et al., 2017).

Several proposals for dose individualization of TKIs have been made, based on PK and PD observations. Therapeutic drug monitoring/management (TDM) of drug plasma concentration appears an obvious choice given the large variability in drug exposure and the apparent exposure–response relationships (Verheijen et al., 2017). The drug exposure may be the area under the curve (AUC) or the pre-dose concentration at steady-state (Cτrough). TDM targets have indeed been suggested for TKIs in previous exposure–response studies. Dose individualization based on PD-biomarkers, such as soluble biomarkers or changes in blood pressure and neutrophil count, has also been proposed for several TKIs (Hansson et al., 2013a; Hansson et al., 2013b; Tsuchiya et al., 2015; Sabanathan et al., 2017; Schmидinger et al., 2017). Challenges in the evaluation of circulating biomarkers reside in (i) the time-related changes (i.e. kinetics) of biomarkers, (ii) their inter- and intra-individual variability and (iii) the limited patient and funding resources compared to the large number of drugs to be investigated (Buyse et al., 2010; Almufti et al., 2014).

In parallel, modification of dosing schedules has been explored for TKIs (i.e. continuous daily dosing versus 4 weeks on, 2 weeks off, Houk et al., 2008; Nozawa and Uemura, 2012). More recently, intermittent administration (i.e. weekly or biweekly) of high-dose TKIs have been proposed (Rovithi et al., 2014; Rovithi et al., 2016). This type of schedule results in higher peak plasma concentrations, and is proposed to give rise to enhanced efficacy, possibly due to a second mechanism of action. It is however not clear what impact this type of schedule has on various adverse effects (AEs) and overall survival (OS). Moreover, biomarker target values might not hold across drug schedules, such as continuous and intermittent, given the various concentration-time profiles different regimens can result in.

In addition to potential changes in clinical outcome, monetary aspects related to dose and schedule adjustments are important to take into consideration (Centanni et al., 2019). In the interest of managing healthcare costs, it has become acknowledged that alternative dosing regimens could result in lower prescribing costs, whilst maintaining equivalent efficacy (Ratain et al., 2019). Cost-effectiveness analysis represents a separate tool that highlights the role of each dose and schedule adjustment from both a clinical and economic perspective (Hughes and Walley, 2001). Typically, however, such analyses employ relatively empirical (i.e. Markov) models to reproduce transition probability between disease stages or between AE grades. Because of this feature, they may lack the flexibility required to draw inferences related to cost-effectiveness of various drug schedules and individualization approaches.

The question of the best treatment schedule and individualization options for TKIs remains, and is complex to sort out, given the different dynamics of each biomarker and AEs. As an illustrative example, diastolic blood pressure was found to have low accuracy in guiding dose adaptation of sunitinib in gastro-intestinal stromal tumors (GIST) (Centanni et al., 2018). Furthermore, there are several AEs variables that need to be taken into consideration as they may hamper dose increments. This in turn give rise to marginal treatment benefits that would require large patient populations to reach adequate power in prospective clinical trials (Goulooze et al., 2016; Centanni et al., 2018). Given the fact that there are over 25 TKIs available, and most are approved for several treatment indications, the number of combinations to be explored would be extensive and require a large amount of monetary and patient resources.

PK/PD modeling and simulation represents a means to analyze available information in order to quantify the added value and dynamics of each dose individualization rationale and dosing schedule in a rapid, ethical and low-cost manner (Almufti et al., 2014; Bender et al., 2015; Standing, 2017; Centanni et al., 2019). Model-based PK/PD frameworks can facilitate schedule selection in the presence of toxicity, disease, and survival models that can emulate a compromise between safety and efficacy under each treatment schedule and biomarker in a structured way (Jänne et al., 2016; Centanni et al., 2018). Moreover, mechanistic PK/PD frameworks possess a more biological character, which aside from potential inaccuracies related to the models, would provide more reliable inferences when analyzing the cost-effectiveness of alternative dosing strategies (van Hasselt et al., 2015). Previous model-based simulations have demonstrated that at the current suggested threshold values, the sunitinib dose in GIST can best be individualized by pharmacodynamic biomarkers, such as neutrophil count or soluble vascular-endothelial growth factor (VEGFR)-3, as opposed to blood pressure and drug concentrations (Centanni et al., 2018).

Given the long treatment period of TKIs such as sunitinib, and the potential impact of AEs on daily life (fatigue, hand-foot syndrome [HFS]), further research on sunitinib in GIST is of importance to optimize the biomarker threshold values for dose adaptation. In addition, the existing model-based framework of axitinib for metastatic renal cell carcinoma (mRCC) can be adapted for similar purposes (Schindler et al., 2017). The following work utilizes the existing sunitinib and axitinib frameworks to evaluate the efficacy, safety, and economical aspects of the abovementioned dosing options. Current clinical protocols and hospital expenses are integrated to emulate real-life scenarios and maximize the direct clinical applicability of the results. In the first section of the work, labeled recommended dosages of axitinib and sunitinib are compared to high-dose intermittent schedules for development of treatment-related AEs and OS. Secondly, biomarker-based dose adjustments are performed to evaluate potential differences in clinical outcomes and thirdly the predictive accuracy of model-based biomarker adjustments is compared to that of regular sample-based methods, wherein dose is directly adjusted according to the measured biomarker plasma levels. Finally, a cost-effectiveness analysis of sunitinib is performed for the available biomarker-based dose interventions.
METHODS

Setting Up the Sunitinib and Axitinib Framework

Two separate sunitinib simulation frameworks were built using a collection of previously created models (Figure 1A) (Supplementary Material 1) (Hansson et al., 2013a; Hansson et al., 2013b). None of the model parameters were re-estimated. Five AE models were available to describe the development of diastolic hypertension, neutropenia, thrombocytopenia, fatigue, and HFS under sunitinib therapy. Here, thrombocytopenia and diastolic hypertensive were driven by sunitinib AUC, whereas remaining AEs were affected by changes in sVEGFR-3. A PK model of sunitinib and its active metabolite (SU12662) was selected to simulate changes in drug concentration over time (Yu et al., 2015). The time-courses of sVEGFR-3 and circulating levels of soluble KIT (sKIT) were described by two indirect response models. Longitudinal tumor size was reproduced with a tumor growth inhibition (TGI) model (Claret et al., 2009), affected by changes in sVEGFR-3, sKIT, and sunitinib AUC. Two OS models were divided over the simulation frameworks; one driven by changes in diastolic blood pressure ($D_{dBP}$), scaled absolute change in neutrophil count (ANC), and baseline tumor size (Eqs. 1 and 2), and the second model was steered by relative changes in sVEGFR-3 ($\Delta sVEGFR-3$) and baseline tumor size (Eq. 3). In the interest of clarity, dose adjustments based on the first (toxicity-related) framework will be referred to as toxicity-adjusted dosing (TAD), whereas dose-adjustments based on the second framework (sVEGFR-3-based) will be referred to as biomarker-based dosing (BAD). Lastly, a dropout model was added to both frameworks to account for the tendency of individuals with poorer responses to discontinue treatment.

In a similar manner, the axitinib simulation framework was built on a collection of previous models, using originally estimated parameters (Figure 1B) (Supplementary Material 2) (Schindler et al., 2017). An axitinib PK model was included to simulate drug serum concentration over time (Garrett et al., 2014). We assumed axitinib was taken orally in the XLI formulation, which is the marketed formulation according to the label information. One AE model was available to describe the development of diastolic blood pressure under treatment. Changes in sVEGFR-3 were captured by an indirect response model. Tumor size (sum of longest diameters, SLD) was described by a growth model, affected by changes in sVEGFR-3. OS, driven by tumor size and progression, was simulated by means of a time-to-event model. Lastly, a dropout model, driven by axitinib exposure, tumor progression, and tumor size, was added to simulate patients who discontinued treatment.

Both the sunitinib and axitinib frameworks were implemented in the R-based package mrgsolve (version 0.8.10, Elmokadem et al., 2019), together with the parameter estimates. Visual audits were performed to compare the mrgsolve to NONMEM (Beal et al., 1989) simulation output to assure accurate translation of the models.

Virtual Populations

In order to achieve the simulations, datasets containing 1,000 virtual individuals with metastatic and/or unresectable GIST, or mRCC, were generated using the dmutate (version 0.1.2) and dplyr (version 0.7.4) R packages. Individual covariates were created by sampling from a defined population distribution (Hansson et al., 2013a; Hansson et al., 2013b; Garrett et al., 2014; Yu et al., 2015; Schindler et al., 2017). Weight was assumed to follow a normal distribution with a mean of 73.5 kg and a standard deviation of 18.7 kg (76.7+/−11.6 kg, for axitinib), truncated between 36 and 185 kg. Baseline tumor size followed a lognormal distribution, with a mean of 182.7 mm and a standard deviation of 134.2 mm, truncated between 29 and 822 mm.

\[
\Delta sVEGFR-3(t) = \frac{sVEGFR-3(t) - BASE_{sVEGFR-3}}{BASE_{sVEGFR-3}}
\]

\[
ANC(t) = ANC(t) - 5
\]

\[
\Delta dBP(t) = \frac{dBP(t) - BASE_{dBP}}{BASE_{dBP}}
\]
Baseline Simulations
Following current guidelines, initial simulations of sunitinib and axitinib were performed using a fixed dose at 37.5 mg daily or 5 mg twice per day (b.i.d.), respectively [Pfizer, Inc. SUTENT® (Sunitinib malate), 2006; Pfizer, Inc. INLYTA® (Axitinib), 2012]. No dose escalations were allowed in the base simulations. Dose reductions were allowed in the event of unacceptable toxicity (≥Grade 3, or ≥Grade 2 for fatigue and HFS), following National Cancer Institute Common Toxicity Criteria (CTCAE) v3.0. To replicate a clinical scenario, monitoring of ANC, dBP, and platelet count occurred solely on days 15, 29, 43, 57, 85, 113, and one every 12 weeks thereafter. Fatigue and HFS were assumed to be spontaneously reported side effects that were monitored on a daily basis.

Dose adjustments were simulated with a discrete number of possible sunitinib and axitinib doses. Available sunitinib doses ranged from 0 to 75 mg of sunitinib once daily, in 12.5 mg increments. For axitinib the available doses were 0, 2, 5, 7, or 10 mg b.i.d. In the event of primary Grade 3 (or Grade 2 for HFS and fatigue) toxicity, the drugs were withheld until ≤ Grade 1 toxicity. Hereafter, the drugs were resumed at the initial dose. In case of repeated, i.e. > 1 occurrence of Grade 3 (or Grade 2 for HFS and fatigue), or severe Grade 4 (or Grade 3 for HFS and fatigue) toxicity, the drugs were withheld until ≤ Grade 1 toxicity, and thereafter resumed at one dose level reduction.

Simulation of High-Dose Treatment
The influence of intermittent high-dose sunitinib or axitinib was evaluated by comparing once weekly (QW), and once every 2 weeks (Q2W) administration, to standard continuous dosing. For these simulations, the available pool of sunitinib doses was extended with doses of 100–700 mg in 100 mg increments. The initial sunitinib dose was fixed to 300 mg for QW and 700 mg for Q2W, with possible dose reduction in the event of AEs (Rovithi et al., 2016). The pool of available axitinib doses was extended with doses of 10–80 mg in 10 mg increments, and 80–140 mg in 20 mg increments, for QW and Q2W dosing regimens, respectively. Axitinib starting doses were at 70 mg for QW, or 140 mg for Q2W, with possible dose reductions in case of AEs.

Identification of Biomarker Cut-Off Points
Different threshold values for each biomarker were evaluated in order to divide biomarker findings into “adjustment” and “no adjustment” dosing decisions. The influence of each biomarker cut-off point was simulated (n = 1,000) to predict population survival and the development of AEs. Individual predicted values of the biomarker (“IPRED”) were used to evaluate the direct effect of dose interventions, i.e. the residual error resulting in a difference between the clinically measured value (“Y”) from the “true” value IPRED, was omitted. Threshold values with similar OS outcomes across TDM, BAD, and TAD were selected in order to ensure that the biomarker-based dose-adjustment approaches could be compared with respect to the development of AEs and cost-effectiveness. If differing OS values would be utilized, the difference between biomarkers could otherwise be caused by discrepancies in biomarker cut-off values. The predicted OS when using the sunitinib TDM threshold was used as reference to ensure equal threshold selections also in the BAD and TAD simulation frameworks.

Cut-off values between 0 ng/ml and 80 ng/ml (sunitinib plus SU12662, or axitinib) were evaluated for TDM of $C_{\text{trough}}$. Since most axitinib exposure-response analyses have focused on the relationship between AUC and response (Verheijen et al., 2017), the correlation between axitinib AUC and $C_{\text{trough}}$ (IPRED) was firstly evaluated using Pearson’s correlation coefficient to warrant the adaptation of $C_{\text{trough}}$ in TDM scenarios.

For TAD based on neutrophils, values between 1 × 109 cells/L and 5 × 109 cells/L were evaluated as cut-offs. Similarly, for BAD based on sVEGFR-3 a range between 5 and 80%, in addition to 0%, decrease from individual sVEGFR-3 baseline was assessed. For TAD based on dBP, cut-off values in the range 2.5 to 40%, in addition to 0%, increase from individual baseline values were evaluated. Based on the biomarker threshold-survival plots, cut-off values were selected for each biomarker that constrained the total number of developed AEs and resulted in equal survival for TDM, TAD, and BAD.

Model-Based Forecasts
The accuracy of clinical sample measurements (“Y”) was evaluated by taking the ratio between the observed change measured in clinical samples at week 8 (BIOMARKER_Y) and the true biomarker change (e.g. Eqs. 1–3) at week 8 (BIOMARKERIPREDtrue) (Eq. 4). The accuracy of model-based estimations of biomarker changes (BIOMARKERIPREDtrue) was evaluated under different monitoring durations and sampling frequencies. The simulation-estimation procedure was performed using the proseval command (prospective evaluation; Pearl-speaks-NONMEM [PsN]) in NONMEM (version 7.3.0) (Beal et al., 1989; Keizer et al., 2013). The function enables automated performance evaluation of maximum a posteriori prediction (MAP) Bayesian estimations based on varying amount of information, i.e. using an increasing number of data points per individual (Abrantes et al., 2019). Based on the dataset of n observations per individual, the command initiates estimations based on the first observation in each individual and runs successive estimations adding one additional observation in each round (in the order of sampling time), until the final number of n observations is reached. As such, the proseval command facilitates the process of determining the influence of observation numbers on individual biomarker value estimation.

In order to determine the influence of monitoring durations and frequencies on the model-based estimations, daily BIOMARKER_Y values were simulated for ANC, dBP, sVEGFR-3, and drug plasma concentration (n = 1,000) using the “true” individual values IPREDtrue with an added random residual error. Datasets containing different numbers of biomarker measurements were generated through changes in (1) monitoring duration (week 0 to week 8) and (2) sampling frequencies (daily, weekly, and biweekly). Based on these outputs, individual parameters were obtained (IPREDest) within the corresponding population pharmacokinetic

---

1 https://github.com/UPPharmacometrics/PsNI/releases/proseval_userguide.pdf
and pharmacodynamic models, using maximum-a-posteriori estimation. The final BIOMARKER_{IPREDtrue} were simulated using the IPRED_{true} values and compared to the individual BIOMARKER_{IPREDest} values simulated with IPRED_{est}. Accuracy of model estimations were computed by taking the ratio of the forecasted estimated biomarker changes (BIOMARKER_{IPREDest}) and true biomarker changes (BIOMARKER_{IPREDtrue}) from baseline at week 8 (Eq. 5).

\[
\text{Accuracy}_{\text{clinical,emp}} = \frac{\text{BIOMARKER}_{Y} + \text{BIOMARKER}_{\text{IPREDtrue}}}{\text{BIOMARKER}_{\text{IPREDtrue}}} \quad (4)
\]

\[
\text{Accuracy}_{\text{IPREDest}} = \frac{\text{BIOMARKER}_{\text{IPREDest}} + \text{BIOMARKER}_{\text{IPREDtrue}}}{\text{BIOMARKER}_{\text{IPREDtrue}}} \quad (5)
\]

The capacity of each method to give rise to correct dose recommendations was further summarized by means of its sensitivity and specificity to inform correct dosing decisions, using the determined biomarker cut-off values. Here sensitivity represents the percentage of times dose increments were recommended at \( t = 1,344 \) h (e.g. day 57/week 8) according to the estimated (BIOMARKER_{IPREDest}) or measured (BIOMARKER_{Y}) changes in biomarker value, compared to the number of times that the “true” biomarker change (BIOMARKER_{IPREDtrue}) would require dose increase. Specificity represents the percentage of times dose increments were not recommended at week 8 according to the estimated (BIOMARKER_{IPREDest}) or measured (BIOMARKER_{Y}) changes in biomarker value, compared to the number of times that the “true” biomarker change (BIOMARKER_{IPREDtrue}) oblige no dose alternations.

**Simulating Cost-Effectiveness**

Cost-effectiveness of biomarker-based dose individualization was determined via differences in absolute costs and quality adjusted life years (QALYs) between standard and individualized treatments. Information regarding treatment costs and hospital expenses of sunitinib therapy were added to the simulation frameworks to analyze the expenditures, as seen in previous work (van Hasselt et al., 2015; Zuidema et al., 2019). Costs of regular follow-up procedures, including medical visits, laboratory investigation, and radiology tests (following clinical guidelines for treatment of metastatic and unresectable GIST), were gathered from the Dutch Healthcare Authority (NZa\(^2\), accessed 12 October 2019). Expenditures related to the development of AEs were collected from (Mickisch et al., 2009), summarizing the costs of each AE in different countries. Because of the absence of Dutch data, the German costs were employed for calculating expenses related to AEs, as the German healthcare system is most compatible to the Dutch one (Zuidema et al., 2019). Drug-related costs were computed by multiplying the price of each sunitinib dose category by the cumulative number of capsules. Prices of each capsule were collected from the Dutch National Health Care Institute (Zorginstituut Nederland\(^3\), accessed 12 October 2019).

To account for potential differences in health status between the biomarker schedules, cost-effectiveness was calculated using both gained life-years and QALYs. In contrast to gained life-years, QALYs combine information regarding both increase in survival time as well as the quality of health during these years, which may be affected by the presence of AEs or changes in disease status. The number of QALYs for a given individual is commonly calculated by multiplying the total number of life-years-gained by the quality of life in these years, generally scored between 0 (equivalent to death) and 1 (equivalent to perfect health) according to the EuroQol- 5 Dimension (EQ-5D) questionnaires (Whitehead and Ali, 2010). In cost-effectiveness analysis, such scores are known as “utility values.” To quantify the overall health quality score at a given moment, utility values of relevant factors should be taken into consideration. For sunitinib, utility values related to baseline status, disease progression, and the development of AEs were here gathered from literature (Paz-Ares et al., 2008; Beauchemin et al., 2016; Liviu Preda and Galieta Mincă, 2018). Individual utility values (UV\(_i\)) at each time point (UV\(_{ij}\)) were computed by multiplying relevant UV\(_i\)'s under disease and the predicted Grades of AE (0–4) (Eq. 6). Individual QALYs (QALY\(_i\)) were then calculated by dividing the sum of daily UV\(_{ij}\) values by days between the start and end of treatment (t\(_{\text{treatment}}\)), and multiplying by the amount of life-years-gained per individual (LYG\(_i\)) (Eq. 7).

\[
UV_{ij} = UV_{\text{neutropenia}_{ij}} \times UV_{\text{thrombocytopenia}_{ij}} \times UV_{\text{hypertension}_{ij}} \times UV_{\text{fatigue}_{ij}} \times UV_{\text{HFS}_{ij}} \times UV_{\text{disease status}_{ij}}
\]

\[
QALY_{i} = LYG_{i} \times \text{sum}(UV_{ij})/t_{\text{treatment}} \quad (7)
\]

Using the established biomarker threshold values, the influence of each sunitinib biomarker was simulated for a virtual patient population (n = 1,000) over a 5-years time horizon, as recommended for cost-effectiveness analysis (Kim et al., 2017) on gained life-years, QALYs, development of AEs, and treatment costs. Due to that AE models were not available for all common AEs during axitinib treatment, this analysis was not possible for the axitinib framework.

**RESULTS**

**High-Dose Intermittent Administration**

Based on the simulation output, OS of the GIST population at week 102 was predicted to be higher under sunitinib standard CD dosing (37.5 mg), compared to QW (300 mg) and Q2W (700 mg) dosing (Table 1). The frequency of high-grade AEs was predicted to be pronouncedly lower for the high-dose schedules, with a relatively low frequency of neutropenia, HFS, and thrombocytopenia for QW, Q2W dosing, compared to CD dosing. The incidence of fatigue was comparable between the QW, Q2W, and CD dosing schedules. Frequency of diastolic hypertension was substantially higher under QW and QW
dosing, compared to CD dosing. For axitinib, OS of the mRCC population at week 102 was predicted to be higher under standard (5 mg b.i.d.) dosing, compared to QW (70 mg) and Q2W (140 mg) dosing.

### Biomarker Threshold Values

The influence of different biomarker threshold values is visualized against changes in OS (Figure 2). The selected sunitinib TDM cut-off value of 50 ng/ml, resulted in 50.5% OS at 102 weeks. For ANC, a 20% decrease from population baseline (e.g. from 5 to 4 × 10⁹ cells/L) resulted in the same OS rate. For BAD, a cut-off value of 25% decrease in sVEGFR-3 from individual baseline was predicted to equal the threshold of 50 ng/ml sunitinib threshold in terms of OS.

Pearson’s correlation coefficient between axitinib AUC and C_{trough} (IPRED) was found to be 0.679 (p < 0.001, n = 1,000). An axitinib TDM cut-off value of 10 ng/ml was selected, resulting in 67.5% OS at 102 weeks. The same OS rate was predicted for an on average 12.5% increase in dBP and a 30% decrease in sVEGFR-3, from individual baselines.

### Biomarker Forecasts

The accuracy of model-based biomarker forecasts compared to using observed clinical biomarker measurements is depicted in Figure 3. The accuracy plots represent the capacity of each method to predict biomarker value at t = 1,344 h (e.g. day 57/week 8). For most biomarkers, model-based predictions demonstrated higher accuracy than direct clinical measurements, particularly when > 1 measurement informed the Bayesian feedback estimation. It is important to notice that IPRED = 0 was estimated solely on the measurement made at t = 0, which adds to the baseline value of the ANC, dBP, and
sVEGFR-3 biomarkers. For the TDM biomarker, however, this represents a sampling moment antecedent to start of drug administration (e.g. a concentration of 0) and IPRED = 0 is therefore representing the population distribution prior to individual knowledge.

The percentage of correct dose decisions that was made at t = 1,344 (e.g. day 57/week 8) following model-based biomarker forecasts (IPRED) or clinical biomarker measurements (Y) is summarized in Figure 4. Here the sensitivity (e.g. percentage of times the method correctly states dose increase should occur) is demonstrated for forecasts based on biweekly, weekly, and daily measurements (following 8 weeks of measurements), as well as for clinical biomarker samples (following one sample at t = 0 and one sample at t = 1,344 h). Specificity (e.g. percentage of times the method correctly states dose increase should not occur) is depicted in a similar manner. Both sensitivity and specificity vary between different biomarkers and prediction method, e.g. model-based prediction versus clinical biomarker measurements, however, model-based estimations generally improve with an increasing measurement frequency.

Cost-Effectiveness Analysis

Based on our sunitinib model we estimated median OS at 1.71, 1.80, 1.90, and 2.16 years for the fixed, TDM-based, ANC-based, and sVEGFR-3-based dosing, respectively (Table 2). The cost per additional life-year was highest for TDM-based (€115,433), when compared to ANC-based (€71,458) and sVEGFR-3-based (€25,340) dosing. Cost per additional QALY was found to be substantially lower for s-VEGFR-3 (€36,784), compared to TDM-based (€173,150) and ANC-based (€104,438) dosing.

DISCUSSION

Two simulation frameworks were created that allowed for survival and safety analysis of different (I) treatment schedules and (II) biomarker-guided dose individualization options for axitinib in mRCC and sunitinib in GIST. Moreover, (III) model-based dose individualizations were weighted against current sample-based approaches. For sunitinib, where relevant AE models were available, the framework was additionally employed to (IV) assess

---

**FIGURE 3 |** Accuracy of biomarker forecasts (IPRED) versus clinical sample-based methods (Y). Accuracy in the simulated population (n = 1,000) as illustrated by the distribution of model-based biomarker forecasts (BIOMARKERIPREDtrue) or directly measured sample values (Y) under continuous bidaily dosing of axitinib at 5 mg or daily dosing of sunitinib at 37.5 mg. Accuracy = [Predicted (BIOMARKERIPREDtrue) or Measured (Y) biomarker change/Actual biomarker change (BIOMARKERIPREDtrue)]. Each plot represents forecasts accuracies at week 8 under different monitoring durations (0, 2, 4, 6, or 8 weeks; IPRED on x-axis) and sampling frequencies (daily [light], weekly [medium], or biweekly [dark]; color shades). For directly measured clinical samples (Y on x-axis) each plot represents the accuracy to capture the true biomarker change at week 8, using one measurement at baseline and one at week 8. Solid black line means 100% accuracy and represents the state at which predicted biomarker change = actual biomarker change. Dashed horizontal lines represent the range wherein the predicted biomarker change falls within the 80–125% range of the true biomarker change. The ends of each box represent the population 25th and 75th percentiles and whiskers the 2.5th and 97.5th percentiles, respectively. Median is represented by the line that cuts through each box in the plot. BAD, biomarker-based dosing; TAD, toxicity-adjusted dosing (ANC for sunitinib; dBP for axitinib); TDM, therapeutic drug monitoring.
the cost-effectiveness of sunitinib therapy under each biomarker option. The abovementioned approaches serve as a methodological example of how model-based frameworks can be applied to assess schedule adjustments, select biomarkers for dose adaptations, and perform pharmacoeconomical analyses.

Simulations from the axitinib and sunitinib frameworks suggest that QW or Q2W high-dosing at the cumulative weekly standard CD dose might result in lower OS in patients with mRCC and GIST, compared to CD dosing. With the exception of diastolic hypertension under sunitinib, the development of AEs was however lower or similar under QW or Q2W high-dose compared to CD dosing, which complies with clinical findings (Rovithi et al., 2018). A possible explanation for this finding resides in the relationship between the axitinib and sunitinib exposure and the development of AEs, which is directly or indirectly defined by Emax equations, with exception of the sunitinib AUC-dBP relationship that is defined by a linear relationship. The adverse effect caused by drug exposure is consequently expected to asymptote to a maximum effect at higher values, thus the high dose regimens is overall resulting in less toxicity over time given that the longer dosing interval allows for washout. The distinction between the development of AEs will be determined by the
longevity of drug effect; e.g. CD will cause more adverse effects compared to QW or Q2W schedules.

It is important to note that extrapolation outside the original dosing schedule, that the models were built on, result in more uncertain predictions and clinical study data would need to confirm the findings and/or serve the basis for refining the models. Since current models were based on frequent, low-dose schedules of axitinib and sunitinib, they may not capture pleiotropic drug effects that can arise at intermittent, high doses. Although the models appear to capture the development of AEs in comparison to clinical data, it is possible that the linear and E\textsubscript{max} equations that are defined using lower dosages may not accurately represent the relationships between drug exposure and effect that occur at higher dosages. As such, schedule adjustments will often necessitate an iterative endeavor of prediction, data generation, and confirmation (EMA extrapolation framework \textsuperscript{4}, accessed 15 November 2019).

Model-based forecast values of biomarkers on >1 clinical measurement provided higher accuracy than most single direct clinical sample measurement values, particularly for dBP under axitinib (Figure 3). Interestingly, previous studies found model-based estimations to have low predictive accuracy for changes in dBP under sunitinib (Centanni et al., 2018). This inconsistency may be due to the presence of inter-individual variability on two additional parameters (i.e. drug effect and dBP turnover rate) and the existence of a combined additive and proportional error instead of just a proportional error. Extrapolation of biomarkers between treatment indications or members of a drug class should therefore be made with caution. In a similar manner, the influence of model-based estimations and clinical sample measurements on the percentage of correct dosing decisions differs between the different biomarkers and drugs (Figure 4). For axitinib, TDM-guided dosing using a model-based method appears to overestimate the number of dose increases that should occur (i.e. low specificity) under the more sparsely sampled (e.g. biweekly) data, compared to just evaluating the clinical measurement at week 8. This is likely due to the initial relative underestimation in drug concentration that is seen in Figure 3. A similar pattern is observed for BAD of sunitinib, where the sensitivity of the model-based estimations is lower than the clinical biomarker measurement when using sparse sampling (Figure 4), due to initial overestimations of biomarkers changes (Figure 3). For the remaining biomarkers, model-based estimations with sparse data provided either an equal or improved percentage of correct dosing decisions, as compared to clinical sample measurements. Importantly, it is expected that model-based estimations are capable of providing predictions close after treatment initiation, whereas clinical samples can only inform dosing decisions once steady-state has been reached (Svensson et al., 2019).

Limitations related to the use of model-based dose individualization for TKIs mainly involve the complexity and computational power related to the model building process, and consequently the availability of such models and software to have a direct implementation at the point of care (van Beek et al., 2019). In addition, since the current models are based upon rich, but often more selected, datasets derived from clinical trials they might require further adjustments to represent the parameter distributions of the entire patient population (Keizer et al., 2018). Advantages of model-based dose individualization of TKIs reside in the (i) capacity to take previous sample observations and patient-specific characteristics into account, (ii) ability to consider a modified dosing history, (iii) flexibility in when to take measurements and finally (iv) capability to predict future effects following dose adjustments.

Cost-effectiveness analysis of biomarkers suggest that sunitinib dose individualization based on sVEGFR-3 results in highest increase in median OS and lowest costs per QALY. Following current clinical expenses, a median increase of €36784 per QALY was determined. Using the threshold of €680 000 and 500,000 SEK (≈€57,000) per gained QALY, this would entail that sVEGFR-3 based dosing can be regarded as a cost-effective intervention in the Netherlands and Sweden, respectively (Vallejo-Torres et al., 2016). The cost-effectiveness of ANC-based and TDM-based dosing appears low, at €104 438 and €173 150 per gained QALY, respectively. Because biomarker threshold values were not set with the primary purpose of increasing OS, alternative biomarker threshold values may optimize cost-effectiveness of ANC-based and TDM-based dosing. Since the correlation between axitinib AUC and C\textsubscript{t\textsubscript{rough}} is not perfect (r = 0.679), it is recommended to perform an exposure-response analysis to further evaluate the correlation between C\textsubscript{t\textsubscript{rough}} and treatment outcome and determine the relevance of C\textsubscript{t\textsubscript{rough}} to guide axitinib dose individualization.

The main disadvantage of cost-effectiveness analysis using a model framework resides in the initial creation of mechanistic models, which demands a significant investment of time and effort. As such, it might primarily appear less desirable. With increasing necessity for model-based analysis in regulatory drug approval it is, however, expected that the required models will become readily available for future developed compounds (Lee et al., 2011). Due to its mechanistic nature it is additionally expected that the predictions made by a model framework are more biologically plausible and therefore better resemble reality (van Hasselt et al., 2015). Moreover, the flexibility of the current model framework the exploration of alternative dosages requires limited additional effort.

To conclude, a model-based framework that includes PK, AEs, circulating biomarkers, tumor growth dynamics, and their relation to OS can be employed to assess alternative treatment schedules and dose-individualization approaches. As more clinical data is collected, each framework can become further extended to reflect the “real-world” population, and to meet novel clinical needs (Keizer et al., 2018). For models in which patient-specific factors (i.e. “covariates”) determine treatment effect, the impact of subpopulations can be simulated and compared to the overall population. As such, the framework provides a flexible approach to evaluate the safety, efficacy, and cost-effectiveness of treatments. In addition to simulations, future efforts should focus on verifying the added value of model-based dose adjustments.

\textsuperscript{4}https://www.ema.europa.eu/en/documents/presentation/presentation-ema-extrapolation-framework-christoph-male_en.pdf
DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

MC performed the analysis and simulations. Both LF and MC contributed to the final version of the manuscript. LF supervised the project.

REFERENCES

Abrantes, J. A., Solms, A., Garman, D., Nielsen, E. I., Jonsson, S., and Karlsson, M. O. (2019). Bayesian Forecasting Utilizing Bleeding Information to Support Dose Individualization of Factor VIII. CPT Pharm. Syst. Pharmacol. 8, 894–903. doi: 10.1002/pssp.12464

Almuffri, R., Wilbaux, M., Orza, A., Henin, E., Freyger, G., Tod, M., et al. (2014). A critical review of the analytical approaches for circulating tumor biomarker kinetics during treatment. Ann. Oncol. 25, 41–56. doi: 10.1093/annonc/mdt382

Arora, A., and Scholer, E. M. (2005). Role of Tyrosine Kinase Inhibitors in Cancer Therapy. J. Pharmacol. Exp. Ther. 315, 971 LP–979. doi: 10.1124/jpet.105.084145

Beal, S. L., Sheiner, L. B., Boeckmann, A. J., and Bauer, R. J. (1989). NONMEM 7.4 users guide. Available at: https://nonmem.iconplc.com/nonmem743/guides. Revised March 2018. Accessed February 2, 2019.

Beauchemin, C., Letarte, N., Mathurin, K., Yelle, L., and Lachaine, J. (2016). A global economic model to assess the cost-effectiveness of new treatments for advanced breast cancer in Canada. J. Med. Econ. 19, 619–629. doi: 10.3111/13696998.2016.1151431

Bender, B. C., Schindler, E., and Friberg, L. E. (2015). Population pharmacokinetic-pharmacodynamic modelling in oncology: a tool for predicting clinical response. Br. J. Clin. Pharmacol. 79, 56–71. doi: 10.1111/bcp.12258

Buyse, M., Sargent, D. J., Grothey, A., Matheson, A., and de Gramont, A. (2010). Biomarkers and surrogate end points—the challenge of statistical validation. Nat. Rev. Clin. Oncol. 7, 309–317. doi: 10.1038/nrclinonc.2010.43

Centanni, M., Krishnan, S. M., and Friberg, L. E. (2018). “A pharmacometric framework for dose individualisation of sunitinib in GIST [Abstract],” in Population Approach Group Europe (Montreux), 27. Abstract nr 8741.

Centanni, M., Moes, D. J. A. R., Trocóniz, I. F., Ciccioni, J., and van Hasselt, J. G. C. (2019). Clinical Pharmacokinetics and Pharmacodynamics of Immune Checkpoint Inhibitors. Clin. Pharmacokinet. 58 (7), 835–857. doi: 10.1007/s40262-019-00748-2

Claret, L., Girard, P., Hoff, P. M., Van Cutsem, E., Zuidevelde, K. P., Jorga, K., et al. (2009). Model-Based Prediction of Phase III Overall Survival in Colorectal Cancer on the Basis of Phase II Tumor Dynamics. J. Clin. Oncol. 27, 4103–4108. doi: 10.1200/JCO.2008.21.0807

Elmkadem, A., Riggs, M. M., and Baron, K. T. (2019). Quantitative systems pharmacology and physiologically-based pharmacokinetic modeling with mrgsolve: a hands-on tutorial. CPT Pharmacometrics Syst. Pharmacol. 8, 883–893

Garrett, M., Poland, B., Brennan, M., Hee, B., Pithavala, Y. K., and Amantia, M. A. (2014). Population pharmacokinetic analysis of axitinib in healthy volunteers. Br. J. Clin. Pharmacol. 77, 480–492. doi: 10.1111/bcp.12206

Goulouze, S. C., Galetti, P., Boddy, A. V., and Martin, J. H. (2016). Monte Carlo simulations of the clinical benefits from therapeutic drug monitoring of sunitinib in patients with gastrointestinal stromal tumours. Cancer Chemother. Pharmacol. 78, 209–216. doi: 10.1007/s00280-016-3071-1

Hansson, E. K., Amantea, M. A., Westwood, P., Milligan, P. A., Houk, B. E., French, J., et al. (2013a). PKPD Modeling of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as Predictors of Tumor Dynamics and Overall Survival Following Sunitinib Treatment in GIST. CPT Pharm. Syst. Pharmacol. 2, e84–. doi: 10.1038/pssp.2013.61

Hansson, E. K., Ma, G., Amantea, M. A., French, J., Milligan, P. A., Friberg, L. E., et al. (2013b). PKPD Modeling of predictors for adverse effects and overall survival in sunitinib-treated patients with GIST. CPT Pharm. Syst. Pharmacol. 2, e85–. doi: 10.1038/pssp.2013.62

Houk, B., Bello, C., Demetri, G., Michaelson, M., Casali, P., Bukowski, R., et al. (2008). Comparative efficacy of sunitinib administered on an intermittent or a continuous daily dosing schedule in metastatic renal cell carcinoma (mRCC) and gastrointestinal stromal tumor (GIST) patients predicted using population PK approaches. Proceedings of the 99th Annual Meeting of the American Association for Cancer Research April 12-16, 2008 San Diego, CA abstr 5828.

Hughes, D. A., and Walley, T. J. (2001). Economic evaluations during early (Phase II) drug development. Pharmacoeconomics 19, 1069–1077. doi: 10.2165/00019053-200119110-00001

Jänne, P. A., Kim, G., Shaw, A. T., Sridhara, R., Pazdur, R., and McKee, A. E. (2016). Dose Finding of small-molecule oncology drugs: optimization throughout the development life cycle. Clin. Cancer Res. 22, 2613 LP–2617. doi: 10.1158/1078-0432.CCR-15-2643

Keizer, R. J., Karlsson, M. O., and Hooker, A. (2013). Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. CPT Pharm. Syst. Pharmacol. 2, e50–. doi: 10.1038/pssp.2013.24

Keizer, R. J., ter Heine, R., Frymoyer, A., Lesko, L. J., Mangat, R., and Goswami, S. (2018). Model-informed precision dosing at the bedside: scientific challenges and opportunities. CPT Pharm. Syst. Pharmacol. 7, 785–787. doi: 10.1002/pssp.12353

Kim, D., Wilkinson, C., Pope, E., Chambers, J., Cohen, J., and Neumann, P. (2017). The influence of time horizon on results of cost-effectiveness analyses. Expert Rev. Pharmacoecon. Outcomes Res. 17 (6), 615–623. doi: 10.1080/14737167.2017.1331432

Lee, J. Y., Garnett, C. E., Gobburu, J. V. S., Bhattachar, V. A., Brar, S., Earp, J. C., et al. (2011). Impact of pharmacometric analyses on new drug approval and labelling decisions. Clin. Pharmacokinet. 50, 627–635. doi: 10.2165/11593210-00000000-00000

Liviu Preda, A., and Galieta Mincă, D. (2018). Cost-effectiveness analysis of treatment for metastatic renal carcinoma in Romania. J. Med. Life 11, 306–311. doi: 10.21522/jml/2018-0069

Mickisch, G., Gore, M., Escudier, B., Procopio, G., Walzer, S., and Nuijten, M. (2009). Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-α2a compared with sunitinib. Br. J. Cancer 102, 80. doi: 10.1038/sj.bjc.6605417

Nozawa, M., and Uemura, H. (2012). Continuous or intermittent? On the dosing schedule of sunitinib for advanced renal cell carcinoma. Transl. Androl. Urol. 1, 202–203. doi: 10.3978/j.issn.2223-4683.2012.05.02

Paz-Ares, L., García del Muro, X., Grande, E., González, P., Brosa, M., and Díaz, S. (2008). Cost-effectiveness analysis of sunitinib in patients with metastatic and/or unresectable gastrointestinal stroma tumours (GIST) after progression or intolerance with imatinib. Clin. Transl. Oncol. 10, 831–839. doi: 10.1007/s12166-008-0297-3

Pﬁzer, Inc. INLYTA® (Axitinib) (2012). Highlights of Prescribing Information. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/202324s008lbl.pdf. Revised August 2018. Accessed February 2, 2019.

FUNDING

The study was supported by the Swedish Cancer Society. CAN 2017/626.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.00316/full#supplementary-material
Pfizer, Inc. SUTENT® (Sunitinib malate) (2006). Highlights of Prescribing Information. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021938s1317s18lbl.pdf [Accessed February 2, 2019].

Ratain, M. J., Goldstein, D. A., and Lichter, A. S. (2019). Interventional pharmacoeconomics—a new discipline for a cost-constrained environment. JAMA Oncol. 5, 1097–1098. doi: 10.1001/jamaoncol.2019.1341

Rovithi, M., Haas, R. R., Honeywell, R. J., Poel, D., Peters, G. J., Griffioen, A. W., et al. (2016). Alternative scheduling of pulsatile, high dose sunitinib efficiently suppresses tumor growth. J. Exp. Clin. Cancer Res. 35, 138. doi: 10.1186/s13046-016-0411-2

Rovithi, M., Gerritse, S. L., Honeywell, R. J., ten Tije, A. J., Ruijter, R., Peters, G. J., et al. (2018). Phase I dose-escalation study of once weekly or once every two weeks administration of high-dose sunitinib in patients with refractory solid tumors. J. Clin. Oncol. 37, 411–418. doi: 10.1200/JCO.18.00725

Sabathan, D., Zhang, A., Fox, P., Coulter, S., Gebiki, V., Balakrishnar, B., et al. (2017). Dose individualization of sunitinib in metastatic renal cell cancer: toxicity-adjusted dose or therapeutic drug monitoring. Cancer Chemother. Pharmacol. 80, 385–393. doi: 10.1007/s00280-017-3362-1

Schindler, E., Amantea, M. A., Karlsson, M. O., and Friberg, L. E. (2017). A pharmacometric framework for axitinib exposure, efficacy, and safety in metastatic renal cell carcinoma patients. CPT Pharm. Syst. Pharmacol. 6, 373–382. doi: 10.1002/psp4.12193

Schmidinger, M., Danesi, R., Jones, R., McDermott, R., Pyle, L., Rini, B., et al. (2017). Individualized dosing with axitinib: rationale and practical guidance. Future Oncol. 14, 861–875. doi: 10.2217/fon-2017-0455

Standing, J. F. (2017). Understanding and applying pharmacometric modelling and simulation in clinical practice and research. Br. J. Clin. Pharmacol. 83, 247–254. doi: 10.1111/bcp.13119

Svensson, R. J., Niward, K., Davies Forsman, L., Bruchfeld, J., Paues, J., Eliasson, E., et al. (2019). Individualised dosing algorithm and personalised treatment of high-dose rifampicin for tuberculosis. Br. J. Clin. Pharmacol. 85, 2341–2350. doi: 10.1111/bcp.14048

Tsuchiya, N., Igarashi, R., Suzuki-Homma, N., Fujiyama, N., Narita, S., Inoue, T., et al. (2015). Association of pharmacokinetics of axitinib with treatment outcome and adverse events in advanced renal cell carcinoma patients. J. Clin. Oncol. 33, 506. doi: 10.1200/jco.2015.33.7_suppl.506

Vallejo-Torres, L., García-Lorenzo, B., Castilla, I., Valcared-Nazco, C., García-Pérez, L., Linertová, R., et al. (2016). On the estimation of the cost-effectiveness threshold: Why, What, How? Value Heal. 19, 558–566. doi: 10.1016/j.jval.2016.02.020

van Beek, S. W., ter Heine, R., Keizer, R. J., Magis-Escura, C., Aarnoutse, R. E., and Svensson, E. M. (2019). Personalized tuberculosis treatment through model-informed dosing of rifampicin. Clin. Pharmacokinet. 58, 815–826. doi: 10.1007/s40262-018-00732-2

van Hasselt, J. G. C., Gupta, A., Hussein, Z., Beijnen, J. H., Schellens, J. H. M., and Huitema, A. D. R. (2015). Integrated simulation framework for toxicity, dose intensity, disease progression, and cost effectiveness for castration-resistant prostate cancer treatment with eribulin. CPT Pharm. Syst. Pharmacol. 4, 374–385. doi: 10.1002/psp4.48

Verheijen, R. B., Yu, H., Schellens, J. H. M., Beijnen, J. H., Steeghs, N., and Huitema, A. D. R. (2017). Practical recommendations for therapeutic drug monitoring of kinase inhibitors in oncology. Clin. Pharmacol. Ther. 102 (5), 765–776. doi: 10.1002/cpt.787

Whitehead, S. J., and Ali, S. (2010). Health outcomes in economic evaluation: the QALY and utilities. Br. Med. Bull. 96, 5–21. doi: 10.1093/bmb/dlt033

Yu, H., Steeghs, N., Kloot, J. S. L., De Wit, D., Van Hasselt, J. G. C., Van Erp, N. P., et al. (2015). Integrated semi-physiological pharmacokinetic model for both sunitinib and its active metabolite SU12662. Br. J. Clin. Pharmacol. 79, 809–819. doi: 10.1111/bcp.12550

Zuidema, S., Desar, I. M. E., Erp, N. P., and Kievit, W. (2019). Optimizing the dose in patients treated with imatinib as first line treatment for gastrointestinal stromal tumours: a cost-effectiveness study. Br. J. Clin. Pharmacol. 85, 1994–2001. doi: 10.1111/bcp.13990

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Centanni and Friberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.