Model-based Dose Individualization of Sunitinib in Gastrointestinal Stromal Tumors

Maddalena Centanni, Sreenath M. Krishnan, and Lena E. Friberg

ABSTRACT

Purpose: Various biomarkers have been proposed for sunitinib therapy in gastrointestinal stromal tumor (GIST). However, the lack of "real-life" comparative studies hampers the selection of the most appropriate one. We, therefore, set up a pharmacometric simulation framework to compare each proposed biomarker.

Experimental Design: Models describing relations between sunitinib exposure, adverse events (hand–foot syndrome, fatigue, hypertension, and neutropenia), soluble VEGFR (sVEGFR)-3, and overall survival (OS) were connected to evaluate the differences in survival and adverse events under different dosing algorithms. Various fixed dosing regimens [4/2 (weeks on/weeks off) or 2/1 (50 mg), and continuous daily dosing (37.5 mg)] and individualization approaches [concentration-adjusted dosing (CAD), toxicity-adjusted dosing (TAD), and sVEGFR-3–adjusted dosing (VAD)] were explored following earlier suggested blood sampling schedules and dose-reduction criteria. Model-based forecasts of biomarker changes were evaluated for predictive accuracy and the advantage of a model-based dosing algorithm was evaluated for clinical implementation.

Results: The continuous daily dosing regimen was predicted to result in the longest survival. TAD (24.5 months) and VAD (25.5 months) increased median OS as compared with a fixed dose schedule (19.9 and 21.5 months, respectively) and CAD (19.7 and 21.3 months, respectively), without markedly raising the risk of intolerable toxicities. Changes in neutrophil count and sVEGFR-3 were accurately forecasted in the majority of subjects (>65%), based on biweekly blood sampling.

Conclusions: Dose adjustments based on the pharmacodynamic biomarkers neutrophil count and sVEGFR-3 can increase OS while retaining drug safety. Future efforts could explore the possibility of incorporating a model-based dose approach in clinical practice to increase dosing accuracy for these biomarkers.

Introduction

Sunitinib is an oral tyrosine kinase inhibitor that is currently deployed for the treatment of imatinib-resistant/intolerant gastrointestinal stromal tumors (GIST), pancreatic neuroendocrine tumors, and metastatic renal cell carcinoma (1). Previous findings have highlighted the strong dose-response relationship between sunitinib and overall survival (OS) in GIST (2), suggesting that dose increments could maximize treatment benefit. Despite the expected improvement in treatment efficacy, sunitinib dose has additionally been associated with a wide range of treatment-related adverse events (AE), such as diarrhea, fatigue, hand-foot syndrome (HFS), hypertension, and myelosuppression (3). The development of such toxicities is problematic, as these not only reduce quality-of-life but also instigate dose reductions and/or delays. Given the strong dose-efficacy relationship, unnecessarily large dose reductions can markedly jeopardize chances of survival. Optimization challenges thus remain to balance maximum therapeutic efficacy in GIST, against the risk of severe adversity.

To improve clinical outcome, alternative dosing schedules of sunitinib for GIST have been proposed. The optimal regimen remains uncertain due to the scarcity of comparative studies between the proposed dosing schedules regarding drug safety and efficacy. Current labels recommend a fixed dose of 50 mg daily for 4 consecutive weeks, followed by a two week off period (4/2 schedule; ref. 1). However, continuous daily dosing of sunitinib at 37.5 mg (CDD schedule) has been associated with an equivalent safety profile (4), whereas daily administration of 50 mg sunitinib for 2 weeks followed by one week off treatment (2/1 schedule) was shown to result in reduced treatment-related AEs (5). On top of this, sunitinib exhibits substantial interindividual variation in exposure, as well as in susceptibility of efficacy and safety endpoints. Individual dosing optimization may thus be motivated (2, 6).

In light of these present issues, dose individualization strategies have been brought forward relying on concentration-adjusted dosing (CAD; refs. 6–9), toxicity-adjusted dosing (TAD; refs. 7, 10–12), or soluble VEGFR (sVEGFR-3)–adjusted dosing (VAD; ref. 13). CAD allows for treatment adjustments based on individual drug exposure patterns and is typically performed on the sum of the total trough concentrations at steady state ($C_{ss,min}$) of the parent drug sunitinib and its active metabolite SU12662 (8, 14). On the basis of previous preclinical (15–17) and clinical (18, 19) studies, a $C_{ss,min}$ acceptable range of 50–100 ng/mL (4/2 and 2/1 schedule) has been suggested to ensure sufficient efficacy and safety (20), which was extrapolated to a $C_{ss,min}$ of 37.5–75 ng/mL for the CDD schedule (20). TAD, on the other hand, allows for treatment adjustments based on individual toxicity profiles (7). This enables patients with high tolerability profiles to increase dose and/or maintain high exposure, even above the recommended $C_{ss,min}$ limit, and profit from maximum treatment efficacy. TAD is of particular relevance for sunitinib, for which AEs such as hypertension and neutropenia have been found to serve as surrogate biomarkers of efficacy and better predictors than changes in tumor size (10–12). Finally, sVEGFR-3 has been related to both AEs of sunitinib and OS in GIST and can therefore also provide viable guidance for dose individualization (13).

Overall, model-based optimization of dose and regimen constitutes an efficient approach to balance sunitinib efficacy against the

Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Corresponding Author: Lena E. Friberg, Uppsala University, Box 591, Uppsala 75124, Sweden. Phone: 461-8471-4685; Fax: 461-8471-4003; E-mail: lena.friberg@farmbio.uu.se

Clin Cancer Res 2020;26:4590–8

doi: 10.1158/1078-0432.CCR-20-0887

Translational Relevance

Sunitinib exerts large interindividual variability in pharmacokinetics (PK) and pharmacodynamics (PD), with clear dose-response relationships. Individualized dosing algorithms can be applied to optimize dose selection with respect to safety and efficacy. PK and PD biomarkers have been suggested to guide dose individualization of sunitinib in gastrointestinal stromal tumor. However, the lack of clinical studies encumbers the selection of the most adequate one. As such, there is a need for rapid, economical, and safe methods to evaluate the clinical relevance of the biomarkers that have been brought forward for sunitinib, and to generate novel hypothesis for clinical evaluation. Modeling and simulations have increasingly been used as methods to analyze data related to drug therapy and explore alternative dosing scenarios. In this simulation study, we demonstrate that dose increases based on the PD biomarkers neutrophil count and sVEGFR-3 can increase OS, while retaining an acceptable safety profile.

development of AEs in patients with GIST. Pharmacometric modeling and simulation can be adopted to quantify the efficacy and toxicity of each proposed algorithm (fixed dosing regimens, TAD, VAD, and CAD), to facilitate schedule and biomarker comparison. The relationships between sunitinib exposure and sVEGFR-3 response to adverse effects (neutropenia, hypertension, fatigue, and HFS) and OS have been previously well characterized by PKPD models (10, 13), whereas the PK of sunitinib has been described by compartmental models (14, 21–23). Integration of the available models in a larger framework will provide an overview of the clinically relevant factors and enable interactions between components (24), and simultaneously explore how dose adjustments affect all variables included in the framework. Performance of such comparisons in "real-life patients" would often be difficult or infeasible in clinical practice, and would require a large amount of monetary resources. Our objectives in this study were therefore to (i) create a simulation framework for sunitinib in GIST, (ii) evaluate which fixed dosing schedule provides the longest OS, (iii) explore the earlier suggested dose-individualization algorithms for the biomarkers CAD, TAD, and VAD and, finally, to (iv) determine the accuracy of model-based forecasts of biomarkers for dose individualization of sunitinib in GIST.

Materials and Methods

Patients and study design

The utilized PKPD models were built based on data from four clinical trials (phases I–III), evaluating sunitinib for imatinib-resistant malignant GIST (3, 4, 25, 26). The database consisted of two phase I/II ($n = 36$; $n = 52$), one phase II ($n = 13$), and one phase III ($n = 249$) trials. Sunitinib was administered orally according to one of the 4/2, 2/2, 2/1 (week on/week off) and CDD treatment schedules. Doses ranged from 0 to 75 mg of sunitinib once daily, by 12.5 mg increments. Trials were approved by local ethics committees and performed in accordance with the ethical principles stated in the Declaration of Helsinki. All study participants provided written informed consent prior to trial inclusion. A summary of the number of patients, along with the evaluated dose and regimen in each trial is provided in the Supplementary Materials (Supplementary Table S1; refs. 3, 4, 25, 26).

Setting up the model framework

To create the simulation framework, models were identified from previous studies (10, 13). Four toxicity models were included to capture the clinically relevant AEs, which were fatigue, HFS, diastolic hypertension, and neutropenia (**Fig. 1**). Of these, only hypertension was directly linked to drug exposure, whereas the remaining models were predicted by changes in sVEGFR-3. Prior to the integration process, correlations, in addition to $C_{ss,min}$/sVEGFR-3, between AE parameters were explored using the Laplace Conditional Estimation method in NONMEM version 7.3 (ICON Development Solutions), to evaluate whether patients with severe AEs were also more prone to develop other toxicities (27). Discrimination between models was based on the likelihood-ratio test, wherein a decrease in objective function value of 3.84 or more upon addition of one extra parameter was considered as significant, which corresponds to $P < 0.05$ (28).

Available pharmacokinetic (PK) models of sunitinib and SU12662 were searched and evaluated for (i) biological plausibility, (ii) availability of model code, and (iii) relevance of the represented population. A literature search leveraged five candidate population PK models of sunitinib (14, 21–23). The model by Yu and colleagues was selected, as it was built upon a sufficient amount of individuals (70 patients) and considered the correlation between important PK parameters of sunitinib and SU12662 (14).

Out of the selected group of models, two separate frameworks were created (**Fig. 1**). In the first "AE-related framework," ANC (scaled absolute change in neutrophil count; Eq. A), ΔdBP (relative change from baseline in diastolic blood pressure; Eq. B), and baseline tumor size, were related to the hazard of death, whereas in the second "sVEGFR-3-related framework," ΔsVEGFR-3 (relative change in sVEGFR-3 from baseline) and baseline tumor size were the main predictors of survival (Eq. C):

$$\mathrm{ANC} = \frac{\mathrm{ANC}(t) - 5}{5} \tag{A}$$

$$\Delta\mathrm{dBP} = \frac{\mathrm{dBP}(t) - \mathrm{BASE}_{\mathrm{dBP}}}{\mathrm{BASE}_{\mathrm{dBP}}} \tag{B}$$

$$\Delta\mathrm{sVEGFR\text{-}3} = \frac{\mathrm{sVEGFR\text{-}3}(t) - \mathrm{BASE}_{\mathrm{sVEGFR\text{-}3}}}{\mathrm{BASE}_{\mathrm{sVEGFR\text{-}3}}} \tag{C}$$

Here, $\mathrm{BASE}_{\mathrm{sVEGFR\text{-}3}}$ and $\mathrm{BASE}_{\mathrm{dBP}}$ represent the baseline values of dBP and sVEGFR-3 at day 0, respectively, while ANC(t), dBP(t), and sVEGFR-3(t) represent the simulated sVEGFR-3 values at time (t; e.g., the dependent variables). For the sake of clarity, dose adjustments according to the first (AE-related) framework will be referred to as TAD and those made with the latter framework (sVEGFR-3-related) as VAD. In each framework, a corresponding dropout model accompanied the survival model (10, 13).

The final collection of models was translated into the R-based package mrgsolve (version 0.8.10; ref. 29) along with the parameter estimates. Comparisons of mrgsolve and NONMEM simulation output were performed with the R-based ggplot2 package (version 2.2.1; ref. 30), to verify the implementation.

Evaluation of dosing regimens

To determine a baseline scenario for further dose individualization, the three fixed dosing schedules (4/2, 2/1, and CDD) were simulated and the influence of each dosing strategy was quantified in terms of

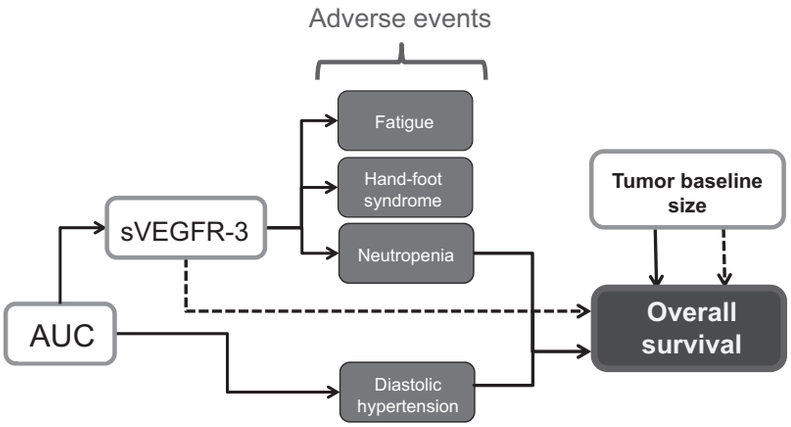

Figure 1.
Model-based framework. Final simulation framework for sunitinib in GIST, adapted from Hansson and colleagues (10, 13). Solid lines represent relationships between the final models. Changes in ANC, HFS, and fatigue were best predicted by relative change in sVEGFR-3 concentration over time, whereas diastolic blood pressure was best predicted by sunitinib AUC. OS was predicted by baseline tumor size, absolute neutrophil count change, and relative blood pressure changes in framework 1 (solid lines). OS was predicted by baseline tumor size and relative sVEGFR-3 change in framework 2 (dashed lines). dBP, diastolic blood pressure; HFS, hand-foot syndrome; sVEGFR-3, soluble VEGFR-3.

survival and development of severe toxicities (≥grade 3 or ≥grade 2 for HFS and fatigue) until week 102, following the NCI Common Toxicity Criteria (CTCAE) v3.0. To replicate a "real-life scenario," measurements of ANC and dBP were taken according to the clinical sample schedule at day 0, 15, 29, 43, 57, 85, 113 and once every 12 weeks thereafter. Fatigue and HFS were assumed to be spontaneously reported side-effects monitored by patients on a daily basis and documented according to an objective scoring method [e.g., FACIT-Fatigue questionnaire in one of the above trials (25)].

Dose individualization strategies

The dosing schedule with the highest OS was selected as a base scenario for the dose individualization strategies. Following the proposed CAD schedule by Lankheet and colleagues, plasma $C_{ss,min}$ of sunitinib and SU12662 were in the simulations assumed to be measured according to clinical sample measurements at day 15, 29, and 57 (9). The daily dose was increased by 12.5 mg in case total plasma concentration was <37.5 ng/mL and if no unacceptable toxicities were present. The daily dose was decreased by 12.5 mg in case the total concentration exceeded 75 ng/mL.

Prior to the TAD and VAD simulations, a sensitivity analysis was carried out to identify the threshold values that guide individual dose adjustments. The effect of various dBP, neutrophil, and sVEGFR-3 threshold points was visualized against changes in AEs, upon which a value was selected that provided maximum dose increase without markedly raising the risk for severe sunitinib toxicities. Simulations with the selected threshold values for dose adaptations were performed according to the same measurement schedule as for CAD, wherein ANC, ΔdBP, and ΔsVEGFR-3 values were determined on day 15, 29, and 57.

For all dosing algorithms (CAD, TAD, and VAD), individual dose adjustments were based on a pool of available doses (0, 12.5, 25, 37.5, 50, 62.5, and 75 mg). The simulated clinical measurement values were represented by the individual model prediction (e.g., IPRED) along with the residual error, to account for clinical measurement errors and random within individual variability.

In all simulations, temporary and permanent dose reductions were simulated to account for the changes that occur following the development of one of the four AEs (i.e., diastolic hypertension, fatigue, HFS, and neutropenia) in real-life clinical practice. The dose was temporarily reduced to 0 mg/day in individuals with first-time grade 3 (or grade 2 for HFS and fatigue) AEs, until ≤grade 1 or resolution, after which dose was resumed at the initial dosage. The dose was temporarily reduced to 0 mg/day in individuals with grade 4 (or grade 3 for HFS and fatigue) AEs, as well as repeated, that is, >1 occurrence of grade 3 (or grade 2 for HFS and fatigue) AEs, until grade 1 or resolution, after which the initial dose was lowered by one level (12.5 mg).

Virtual population

Virtual patient populations consisting of 1,000 individuals were created with the R-based dmutate (version 0.1.2) and dplyr (version 0.7.4) software packages. Values of influential covariates, such as body weight and baseline tumor size were generated to be similar to the original population: (i) a normal distribution of weight with a mean of 73.5 kg and a SD of 18.7 kg, truncated between 36 and 185 kg, and (ii) a log-normal distribution of baseline tumor size with a mean of 182.7 mm and a SD of 134.2 mm, truncated between 29 and 822 mm. Due to absence of known correlations between these covariates, values were randomly sampled from the diagonal elements, without adaptation of off-diagonal ones. Baseline probabilities for the categorical data were set to be the same as for the original patient population (**Fatigue**: grade 0: 0.9538, grade 1: 0.023, grade 2: 0.0198, grade ≥3: 0.0033. **HFS**: grade 0: 1, grade 1: 0, grade 2: 0, grade ≥3: 0).

Statistical analysis

Statistical analyses were performed using the R-based survminer (version 0.4.6) and stats (version 3.5.2) software packages. A Bonferroni-corrected log-rank test was utilized to establish statistically significant differences in OS between dosing scenarios. Differences in the frequency of AEs were determined by means of a χ^2 test (31, 32). Because TAD and VAD scenarios were simulated with their corresponding model frameworks, that is, different OS models, a direct comparison of simulation results between these two approaches was not possible. Instead, CAD and fixed dosing schedule were compared with TAD and VAD separately (i.e., two analyses, each with a fixed dosing regimen and two dose-adjustment arms).

To establish a trial size at which TAD dosing is superior to fixed dosing, we simulated clinical trials wherein half of the patients were assigned to a conventional therapy arm and half received biomarker-based dose individualization. The process was repeated with varying number of trial participants (500, 1,000, and 2,000) at a follow-up time of 102 weeks, to determine the number of study participants required to provide a power of 80%. Power was calculated as the percentage of trials out of the simulated studies in which a statistically significant ($P <$ 0.05) difference in OS was detected, as suggested by previous work of Kowalski and colleagues (33). The simulations accounted for changes in population size over time related to patient death and dropout.

Model-based forecasts

Maximum a posteriori (MAP) estimation was used to calculate individual parameters (empirical Bayes estimates) based on an increasing number of individual biomarker measurements (34). This was achieved by minimizing the MAP objective function using the proseval command in Pearl-speaks-NONMEM (PsN; refs. 35, 36; Eq. D):

$$\mathrm{OBJ}_{\mathrm{MAP}} = \sum_{j=1}^{p} \left[\frac{\hat{P}_j - P_j}{\omega_j}\right]^2 + \sum_{i=1}^{m} \left[\frac{\hat{A}_i - A_i}{\sigma_i}\right]^2 \quad \text{(D)}$$

Here $\hat{P}_j$ represents the typical population estimate, P_j the predicted individual parameter, and ω_j represents the variance of the jth population parameter. A_i symbolizes the observed PD measurement (e.g., neutrophil count), $\hat{A}_i$ the predicted PD measurement, and σ_i the variance of the residual error.

Daily ANC, dBP, and sVEGFR-3 measurements were simulated for 1,000 patients from the corresponding models (10, 13). The true individual values ($\mathrm{IPRED}_{\mathrm{true}}$) based on simulated individual parameters (i.e., no added residual error) were recorded for each variable at the given time points. Expanding biomarker monitoring durations (day 0 up to day 28) and increasing frequencies (daily, weekly, and biweekly) were explored to estimate individual model parameter values. The estimated individual parameters were subsequently used to forecast values of the variables ($\mathrm{IPRED}_{\mathrm{pred}}$) for each individual. The accuracy of the forecasts was assessed by comparing the true IPRED values with the predicted IPRED values for each individual (Eq. E):

$$\mathrm{Accuracy} = \mathrm{IPRED}_{\mathrm{pred}}/\mathrm{IPRED}_{\mathrm{true}} \quad \text{(E)}$$

Forecasts were deemed accurate if the $\mathrm{IPRED}_{\mathrm{pred}}$ value was within 80%–125% of the $\mathrm{IPRED}_{\mathrm{true}}$ value (accuracy = 0.8–1.25).

Results

Comparison of dosing regimens

Simulations with fixed dosing schedules revealed that the more dose-intense CDD schedule performed either statistically or numerically better than the 4/2 and 2/1 schedules in terms of OS ($n = 1{,}000$/arm: $P = 0.026$ and $P = 0.11$, respectively). The median OS was 20.8 months (95% CI, 19.8–21.9, for 1,000 individuals) under the CDD schedule versus 17.7 months (95% CI, 16.8–19.8) under the 4/2 schedule and 18.5 months (95% CI, 17.2–20.4) under the 2/1 schedule. On the basis of this advantage in OS, CDD was selected as the regimen for the dose individualization scenarios. The development of AEs was overall similar across dosing schedules, with the exception of the ≥grade 2 HFS, which was higher under the CDD schedule (4.1%–10.4% for CDD vs. 1.3%–5% for 2/1 and 4/2 dosing: $P < 0.001$ and $P < 0.001$, respectively) and ≥grade 3 neutropenia, which was lower under the 2/1 schedule (0.1%–1.6% for 2/1 vs. 1.0%–4.1% for CDD and 4/2 dosing: $P = 0.13$ and $P = 0.02$, respectively; $n = 1{,}000$/arm).

Simulations of dosing algorithms

On the basis of the sensitivity analysis of various ANC, ΔdBP, and ΔsVEGFR-3 threshold values, limits for dose escalations in TAD and VAD approaches were set to ANC >−0.27, ΔdBP <0.075, and ΔsVEGFR-3 >−0.45. Accordingly, dose titration was performed whenever changes from the baseline value were less than a 45% decrease for sVEGFR-3, 27% decrease for ANC or 7.5% increase for dBP. To prevent excessive AE development in patients with outlying absolute values, dose increments were solely allowed in patients with a dBP <90 mmHg and ANC $>2.0 \times 10^9$/L.

Both TAD and VAD increased OS as compared with fixed dose and CAD-based dose adjustments (**Fig. 2**). The median OS was 24.5 months (95% CI, 21.8–26.8, for 1,000 individuals) in the TAD arm versus 19.9 months (95% CI, 18.3–21.9) in the fixed dosing arm and 19.7 months (95% CI, 18.3–21.7) in the CAD-based dosing arm. Likewise, for the VAD scenario, the median OS was 25.5 months (95% CI, 23.0–27.9) in the VAD arm versus 21.5 months (95% CI, 20.2–23.5) in the fixed dosing arm and 21.3 months (95% CI, 20.2–21.7) in the CAD-based dosing arm.

CAD reduced the incidence of therapy-induced AEs after one treatment cycle as compared with the fixed dosing regimen, but this was found to be statistically significant for neutropenia alone (neutropenia: $P = 0.034$, fatigue: $P = 0.55$, HFS: $P = 0.23$; $n = 1{,}000$/arm; **Fig. 2**). TAD and VAD on the other hand increased the incidence for HFS in comparison with fixed dosing [from 10.8% vs. 13.1% ($P = 0.11$) and 10.8% vs. 13.5% ($P = 0.06$), in cycle 5]. Similarly, TAD and VAD increased the incidence for diastolic hypertension [from 2.1% vs. 2.3% ($P = 0.76$) and 2.1% vs. 2.4%: ($P = 0.65$), in cycle 1] and neutropenia (from 4.4% vs. 4.9% ($P = 0.60$) and 4.4% vs. 5.1%: ($P = 0.46$), in cycle 1] in comparison with fixed dosing. All four dosing algorithms demonstrate reduction in the incidence of AEs over consecutive treatment cycles.

Power analysis

Study power was assessed by means of trial simulation with 500, 1,000, and 2,000 participants, at a follow-up period of 102 weeks. A clinical trial comparing OS for standard-of-care fixed CDD schedule versus TAD needs to include at least 1,000 patients per treatment arm to obtain 80% power at an alpha of 0.05.

Forecast accuracy

The accuracies of ANC, ΔdBP, and ΔsVEGFR-3 forecasts are presented in **Fig. 3**. For all forecasts, the spread of the accuracy values narrowed with increased monitoring durations and monitoring frequencies. Generally, the relative prediction errors were larger for ΔdBP forecasts than for ANC and ΔsVEGFR-3. Given the predefined prerequisite (80%–125%), forecasts of ANC were predicted accurately for 70.7%, ΔdBP for 28.4%, and ΔsVEGFR-3 for 63.8% of the individuals at day 29, after biweekly measurements (day 14 and 28). On the basis of these results, biweekly measurements were regarded to be sufficiently informative for ANC and ΔsVEGFR-3 forecasts. Because the accuracy of ΔdBP forecasts in general was low, further evaluation of ΔdBP-based dose adjustments was based on daily dBP measurements.

Discussion

This study has taken a unique approach by merging drug concentration, multiple AEs, biomarker, and survival models into two comprehensive frameworks to simulate sunitinib treatment in imatinib-resistant/intolerant GIST, with the aim of comparing proposed dose individualization scenarios. All models were successfully combined to represent current clinical practices, such as discrete number of possible sunitinib doses (0–75 mg, by 12.5 mg increments), toxicity-induced dose reductions and defined blood sampling days. Except for the correlation related to drug concentration, no correlations between the AEs were identified. The combined modeling framework enabled us to directly compare all proposed fixed dosing regimens (4/2, 2/1, and CDD) and individualized dosing algorithms (CAD, TAD, and VAD) with respect to four AEs and OS, and to account for clinical limitations, such as sparse blood sampling days and dose adjustments in response to intolerable AEs. As such, our results provide a comprehensive overview of treatment outcomes under the currently proposed sunitinib dosing regimens for GIST.

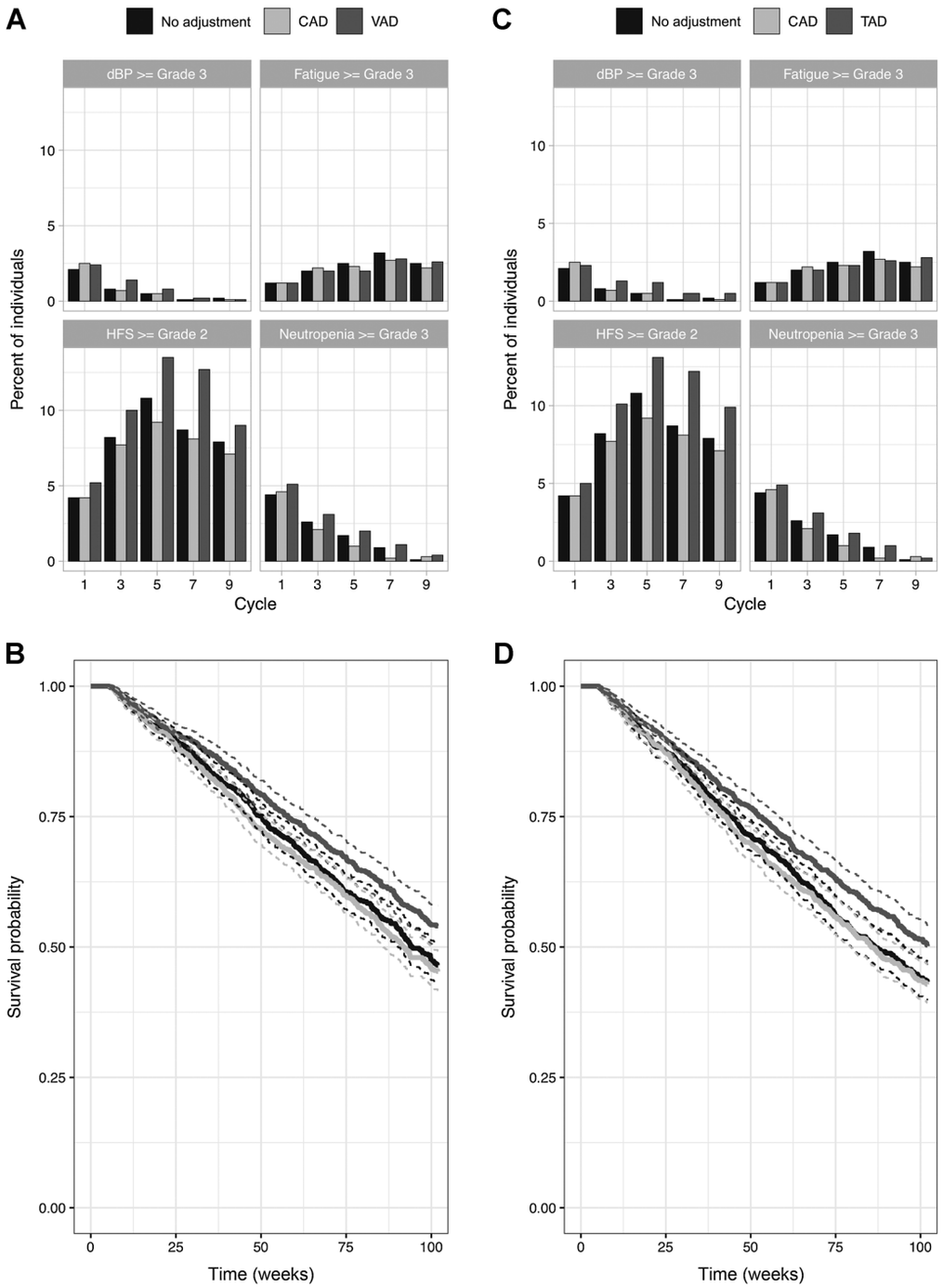

Figure 2.

Safety and efficacy under dose individualization option. **A** and **C,** Development of AE under each (dose individualization) strategy $n = 1{,}000$. One cycle represents 6 weeks of treatment. **B** and **D,** Survival under each treatment regimen, solid lines represent population median, dashed lines represent population 2.5th and 97.5th percentiles ($n = 1{,}000$). Time is represented as weeks (week 0–102) following the initiation of treatment. CAD, concentration-adjusted dosing; dBP, diastolic blood pressure; HFS, hand-foot syndrome; TAD, toxicity-adjusted dosing; VAD, sVEGFR-3–adjusted dosing.

A previous simulation study evaluated the differences between the sunitinib 4/2 and 2/1 dosing schedules with respect to neutrophil count and diastolic blood pressure (37). In alignment with these results, we showed that these fixed dosing schedules share a comparable safety profile, with <2.5% difference in the incidence of ≥grade 3 neutropenia and diastolic hypertension between the 4/2 and 2/1 schedule at steady state. The incidence of HFS in the initial five treatment cycles (e.g., week 0–30) is notably higher under the CDD schedule (4.1%–10.4%),

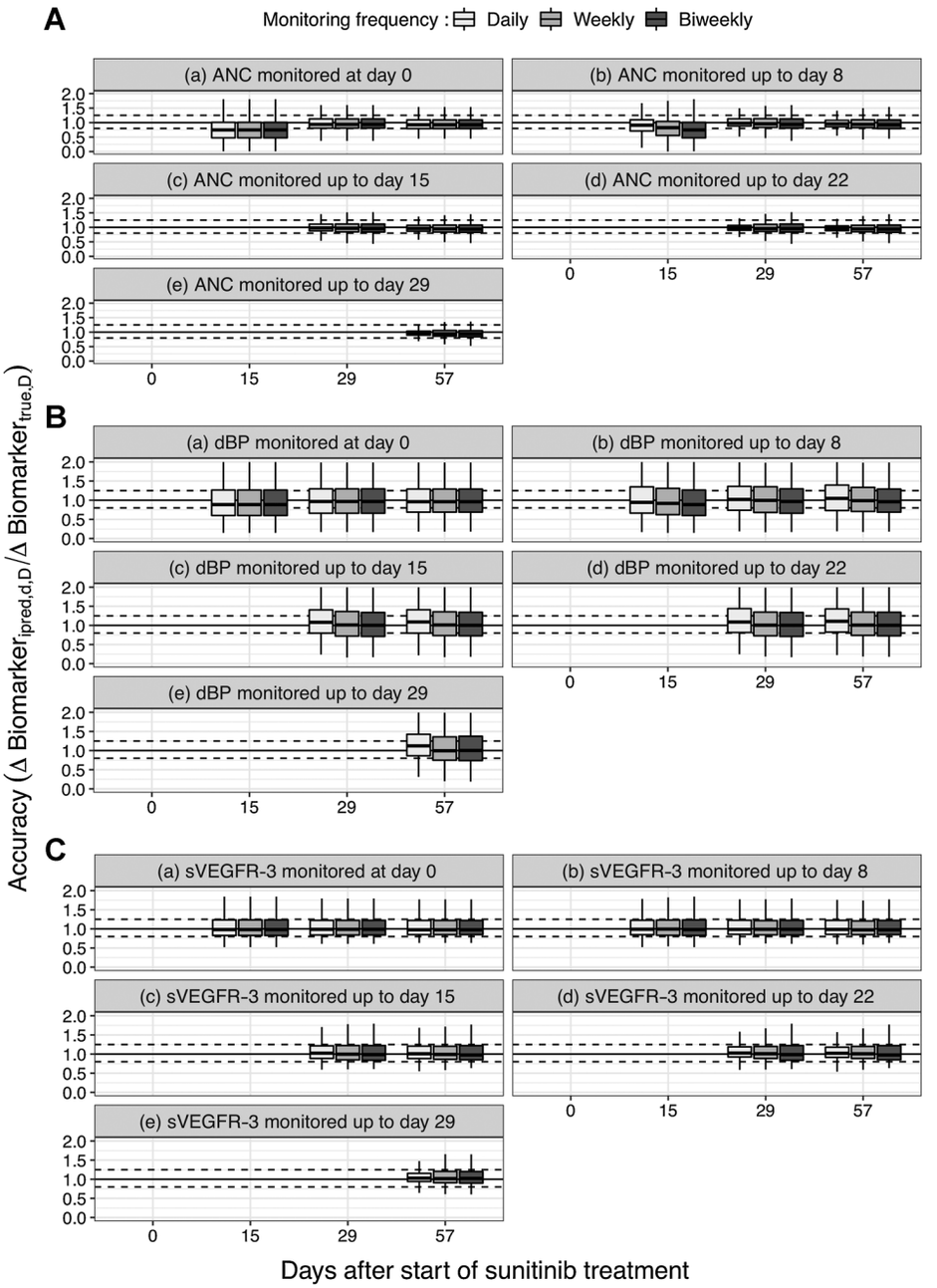

Figure 3.

Model-based biomarker forecasts. **A,** ANC forecasts. **B,** ΔdBP forecasts. **C,** ΔsVEGFR-3 forecasts. Distribution of relative prediction error for scenarios wherein biomarker has been monitored until day 0, 7, 14, 21, and 28 (under sunitinib CDD of 37.5 mg) $n = 1{,}000$. The colours represent a different monitoring frequency of daily (light), weekly (medium), and biweekly (dark) measurements. Ipred, d, D represents the individual predicted biomarker value (without residual error) at day D, provided measurements have been available up until day d, and true, D represents the "true" individual biomarker value (without residual error) at day D. The solid black line represents the value at which the predicted value matches the true value 100%. The dashed horizontal black lines represent the accuracy range wherein the predicted value falls within 80%–125% of the true value. For each boxplot, the vertical lines represent the median value, the hinges correspond to the 25th and 75th percentiles, and the whiskers are the 2.5th and 97.5th percentiles, respectively. ANC, absolute neutrophil count; dBP, diastolic blood pressure; sVEGFR-3, soluble VEGFR-3.

compared with the 2/1 and 4/2 dosing (1.3%–5%). This differs from the findings in two clinical trials, wherein the difference between dosing schedules was less pronounced [5%–14% for CDD (NCT00130897) vs. 9%–12% for 4/2 dosing (NCT00083889); ref. 38]. It should, however, be noted that these findings were derived from two separate populations with diverging distribution in patient-specific characteristics (39, 40), which may have influenced the AE incidences. With the exception of this finding, however, the CDD schedule appeared to provide a good safety versus efficacy profile. Taking this into account, the CDD schedule was selected as baseline option for further simulations.

In the CAD scenario, we found that approximately 20% of the patients required dose increase with a lower threshold concentration of 37.5 mg/L, while in the previous simulation study, dose increments were suggested for 27% of the individuals for the same target (14). However, our simulation study also considered dose restrictions due to toxicity, which limited the possibility for dose increase in some patients, whereas dose decisions within the previous publication were exclusively based upon drug concentration values. In fact, the clinical study by Lankheet and colleagues found that 52% of patients were observed to be below the target of 37.5 ng/mL, but only 17% managed dose escalation without additional ≥grade 3 toxicity (9), suggesting

that the simulated 20% more closely resembles the clinical scenario. Our findings suggest that CAD with the currently suggested acceptable range for CDD of sunitinib does not significantly improve OS, in contrast to what has been previously suggested (6, 8, 9, 20). CAD at days 14 and 28 reduces the incidence of AEs after one treatment cycle (6 weeks; **Fig. 2**), suggesting that even though CAD might not increase therapeutic efficacy, it could improve safety. However, in our simulations the reduction in AEs was only found to be statistically significant for the occurrence of neutropenia. Absence of clear benefit resulting from CAD has been predicted in a previous model-based analysis, wherein the large variability in EC_{50} for efficacy, as well as the safety endpoints, was anticipated to prohibit benefit from therapeutic drug monitoring (TDM; ref. 37). In addition, the current acceptable $C_{ss,min}$ range might be based on a too limited amount of preclinical and clinical knowledge (6, 8, 20). Preclinical IC_{50} values have been recently demonstrate to relate poorly to clinical exposure-response targets (41), and the sunitinib IC_{50} of 50–100 nmol/L (equivalent to 26.6–53.26 ng/mL of sunitinib malate) in several publications (15, 17), were extrapolated to a value of 50–100 of ng/mL (6, 20). Moreover, the motivation for a minimum acceptable $C_{ss,min}$ of 50 ng/mL was based on a dose response analysis (19; $n = 28$), rather than the preferred exposure-response analysis, whereas an exposure-analysis revealed a minimum acceptable $C_{ss,min}$ of 24.6 ng/mL (extrapolated from an AUC of 600 ng*h/mL; ref. 42). The maximum acceptable $C_{ss,min}$ of 100 ng/mL was defined given a higher toxicity incidence at $C_{ss,min} >$ 100 ng/mL (23.1% vs. 75.%; ref. 18; $n = 21$; refs. 6, 8, 20). In fact, a lower toxicity threshold of 60 ng/mL was suggested in a more recent analysis (43). Overall, these findings translate into favoring PD biomarkers over PK ones for dose individualization of sunitinib in GIST.

For the biomarker-based dose titrations, we found that TAD and VAD significantly increased OS in comparison with other dosing strategies. While the incidence in HFS, neutropenia and diastolic hypertension also appeared higher under these regiments on visual inspection, these changes were not found to be statistically significant (**Fig. 2**). Following these results, it would be fruitful to further explore these pharmacodynamic biomarkers in the clinical setting. For TAD, which is based on biomarkers that are readily measured in the clinic, an initial prospective or retrospective cohort study could be conducted to evaluate the association between neutrophil changes and treatment outcomes in a "real-world scenario." Further validation of biomarkers should preferably occur in a larger, randomized clinical trial. To prove clinical improvement in survival under TAD, we found that a large number of participants (1,000/arm) would be required to provide sufficient power at an alpha of 5%. This is partly due to the presence of dose-limiting toxicities, which reduce the possibility for dose escalation, and the relatively small changes in OS expected (44). Dose adjustments could be made on the basis of measurements at week 2, 4, and 8, following the schedule by Lankheet and colleagues, but alternative schedules might be considered to comply with hospital practice, treatment guidelines, and patient preferences. In addition, because the peak of HFS appears to occur at later time points (cycle 5; week 30–36), it is useful to include additional biomarker follow-up points at several months after the dose adjustments. Because we explored a large, but finite, number of biomarker threshold values, further optimization of schedule days and biomarker threshold values could become achieved computationally, by means of multiobjective optimization (45) with logistic constraints (46).

Bayesian forecasts appear sufficiently accurate to identify changes in neutrophil count and sVEGFR-3 (**Fig. 3**). It is expected that over a longer follow-up period, or with the incorporation of additional blood samples, model-based predictions would further improve. Moreover, model-based estimations have the capacity to provide recommendations shortly after treatment initiation, while decisions based on only the clinical sample measurements at hand require steady state to be reached (47). For diastolic blood pressure, forecasts are not sufficiently precise to support model-based dose adjustments and are therefore not recommended (**Fig. 3**). An explanation for this low accuracy could reside in the relatively large residual-unexplained variability (e.g., due to random within-patient variability) of dBP compared with other biomarkers less susceptible to within- and between-day variations, which obscures calculation of individual parameter values. Large residual-unexplained variability may be the reason to why individual dose titration based on blood pressure failed to demonstrate treatment benefit for axitinib (48). On the basis of these results, we recommend dose adjustments in the TAD scenario to be made on the basis of neutrophil count only, without taking dBP into consideration. Absence of dBP is not expected to largely influence the results, because (i) simulations with large residual-unexplained variability still gave good OS and (ii) during the model-build process the time course of neutrophils was the most significant predictor of OS (10). Nevertheless, further investigations to improve the accuracy of dBP as a biomarker, could further support accurate dose adjustments. Because the original model could not account for certain dynamic (e.g., circadian rhythm) and stochastic (e.g., interoccasion variability) sources of dBP variability due to lack of informative data, future efforts could account for these factors to minimize the residual-unexplained variability.

Even though the framework can provide valuable inferences for clinical practice, there are several aspects that might influence extrapolation of our results to clinical implementation. First, our proposed measurement schedule for CDD may not be directly transferable to the 4/2 and 2/1 schedules that exhibit fluctuations in the AE time course over one treatment cycle. CDD has the advantage that biomarkers will approach a steady state, which will facilitate interpretation of clinical measurements, and CDD predictions indicate higher OS. Therefore, the CDD schedule may be in favor over the intermittent schedules. Also, we did not optimize the CAD target window, but rather performed dose adjustments based on the currently recommended threshold values. An increase in maximum $C_{ss,min}$ is expected to provide an increased OS, but may also increase the incidence of AEs. Moreover, it was not explored if a combination of CAD with TAD or VAD would be beneficial. Finally, we did not account for individuals that received antihypertensive treatment prior to sunitinib therapy (GIST: 8%–15%; refs. 49, 50), but as discussed above, dBP does not appear to be an attractive biomarker for dose individualization in the first place.

In summary, a simulation framework was developed on the basis of integration of a PK model, models of four AEs, a model for the sVEGFR-3 biomarker, and OS. The framework was applied to compare outcomes of different dosing schedules, as well as dose-individualization biomarkers, in a realistic and unified manner. TAD (e.g., scaled ANC) and VAD (e.g., sVEGFR-3 concentration) were, based on the proposed threshold values, found to provide the best balance between safety and efficacy of sunitinib in GIST. As such, ANC or sVEGFR-3 might provide a viable guide for dose individualization in the clinic. Neutrophil-based dose adjustments may be preferred, as these are frequently measured in the clinic and will not warrant additional hospital visits or expenses, and it was demonstrated that neutrophil counts could be predicted with reasonable adequacy. The framework could be further applied to fine tune the AE thresholds for

dose titration toward increased efficacy or a safety-favoring scenario, or to explore alternative biomarker sampling schedules. Future efforts could expand the methodologic concept toward other drugs and cancer types. For sunitinib in GIST, we suggest that further studies are conducted to explore the possibility of using neutrophil count and sVEGFR-3–based dose individualization in a "real-world scenario." To our knowledge, this is the first simulation study for dose individualization that includes drug exposure, biomarker concentration, several AEs, and OS.

Disclosure of Potential Conflicts of Interest

L.E. Friberg reports grants from Swedish Cancer Society during the conduct of the study; grants from Genentech; and other from MerckSerono (Commissioned work) outside the submitted work. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

M. Centanni: Methodology, writing-original draft. **S.M. Krishnan:** Methodology. **L.E. Friberg:** Supervision, funding acquisition, writing-original draft.

Acknowledgments

L. Friberg was supported by the Swedish Cancer Society (CAN 2017/626).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 6, 2020; revised May 12, 2020; accepted June 3, 2020; published first June 10, 2020.

References

1. Pfizer, Inc. SUTENT (Sunitinib malate). Highlights of Prescribing Information [Internet]. 2006. [cited 2019 Feb 2]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021938s13s17s18lbl.pdf
2. Houk BE, Bello CL, Poland B, Rosen LS, Demetri GD, Motzer RJ. Relationship between exposure to sunitinib and efficacy and tolerability endpoints in patients with cancer: results of a pharmacokinetic/pharmacodynamic meta-analysis. Cancer Chemother Pharmacol 2010;66:357–71.
3. Demetri GD, van Oosterom AT, Garrett CR, Blackstein ME, Shah MH, Verweij J, et al. Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial. Lancet 2006;368:1329–38.
4. George S, Blay JY, Casali PG, Le Cesne A, Stephenson P, DePrimo SE, et al. Clinical evaluation of continuous daily dosing of sunitinib malate in patients with advanced gastrointestinal stromal tumour after imatinib failure. Eur J Cancer 2009;45:1959–68.
5. Britten CD, Kabbinavar F, Randolph Hecht J, Bello CL, Li J, Baum C, et al. A phase I and pharmacokinetic study of sunitinib administered daily for 2 weeks, followed by a 1-week off period. Cancer Chemother Pharmacol 2008;61:515–24.
6. Westerdijk K, Desar IME, Steeghs N, Graaf WTA, Erp NP, (DPOG) on behalf of the DP, et al. Imatinib, sunitinib and pazopanib: from flatfixed dosing towards a pharmacokinetically guided personalized dose. Br J Clin Pharmacol 2020;86:258–73.
7. Sabanathan D, Zhang A, Fox P, Coulter S, Gebski V, Balakrishnar B, et al. Dose individualization of sunitinib in metastatic renal cell cancer: toxicity-adjusted dose or therapeutic drug monitoring. Cancer Chemother Pharmacol 2017;80:385–93.
8. Verheijen RB, Yu H, Schellens JHM, Beijnen JH, Steeghs N, Huitema ADR. Practical recommendations for therapeutic drug monitoring of kinase inhibitors in oncology. Clin Pharmacol Ther 2017;102:765–76.
9. Lankheet NAG, Kloth JSL, Hooijdonk CGMG, Cirkel GA, Mathijssen RHJ. Pharmacokinetically guided sunitinib dosing : a feasibility study in patients with advanced solid tumours. Br J Cancer 2014;110:2441–9.
10. Hansson EK, Ma G, Amantea MA, French J, Milligan PA, Friberg LE, et al. PKPD modeling of predictors for adverse effects and overall survival in sunitinib-treated patients with GIST. CPT Pharmacometrics Syst Pharmacol 2013;2:e85.
11. George S, Reichardt P, Lechner T, Li S, Cohen DP, Demetri GD. Hypertension as a potential biomarker of efficacy in patients with gastrointestinal stromal tumor treated with sunitinib. Ann Oncol 2012;23:3180–7.
12. Donskov F, von Mehren M, Cams A, George S, Casali RG, Li S, et al. 1138 POSTER neutropenia as a biomarker of sunitinib efficacy in patients (Pts) with gastrointestinal stromal tumour (GIST). Eur J Cancer 2017;47:S135.
13. Hansson EK, Amantea MA, Westwood P, Milligan PA, Houk BE, French J, et al. PKPD modeling of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as predictors of tumor dynamics and overall survival following sunitinib treatment in GIST. CPT Pharmacometrics Syst Pharmacol 2013;2:e84.
14. Yu H, Steeghs N, Kloth JSL, De Wit D, Van Hasselt JGC, Van Erp NP, et al. Integrated semi-physiological pharmacokinetic model for both sunitinib and its active metabolite SU12662. Br J Clin Pharmacol 2015;79:809–19.
15. Abrams TJ, Lee LB, Murray LJ, Pryer NK, Cherrington JM. SU11248 inhibits KIT and platelet-derived growth factor receptor β in preclinical models of human small cell lung cancer. Mol Cancer Ther 2003;2:471–8.
16. Mendel DB, Laird AD, Xin X, Louie SG, Christensen JG, Li G, et al. *In vivo* antitumor activity of SU11248, a novel tyrosine kinase inhibitor targeting vascular endothelial growth factor and platelet-derived growth factor receptors. Clin Cancer Res 2003;9:327–37.
17. Murray LJ, Abrams TJ, Long KR, Ngai TJ, Olson LM, Hong W, et al. SU11248 inhibits tumor growth and CSF-1R-dependent osteolysis in an experimental breast cancer bone metastasis model. Clin Exp Metastasis 2003;20:757–66.
18. Noda S, Otsuji T, Baba M, Yoshida T, Kageyama S, Okamoto K, et al. Assessment of sunitinib-induced toxicities and clinical outcomes based on therapeutic drug monitoring of sunitinib for patients with renal cell carcinoma. Clin Genitourin Cancer 2015;13:350–8.
19. Faivre S, Delbaldo C, Vera K, Robert C, Lozahic S, Lassau N, et al. Safety, pharmacokinetic, and antitumor activity of SU11248, a novel oral multitarget tyrosine kinase inhibitor, in patients with cancer. J Clin Oncol 2006;24:25–35.
20. Yu H, Steeghs N, Nijenhuis CM, Schellens JHM, Beijnen JH, Huitema ADR. Practical guidelines for therapeutic drug monitoring of anticancer tyrosine kinase inhibitors: focus on the pharmacokinetic targets. Clin Pharmacokinet 2014;53:305–25.
21. Houk BE, Bello CL, Kang D, Amantea M. A population pharmacokinetic meta-analysis of sunitinib malate (SU11248) and its primary metabolite (SU12662) in healthy volunteers and oncology patients. Clin Cancer Res 2009;15:2497–506.
22. Lindauer A, Di Gion P, Kanefendt F, Tomalik-Scharte D, Kinzig M, Rodamer M, et al. Pharmacokinetic/Pharmacodynamic modeling of biomarker response to sunitinib in healthy volunteers. Clin Pharmacol Ther 2010;87:601–8.
23. Diekstra MH, Fritsch A, Kanefendt F, Swen JJ, Djar M, Sörgel F, et al. Population modeling integrating pharmacokinetics, pharmacodynamics, pharmacogenetics, and clinical outcome in patients with sunitinib-treated cancer. CPT Pharmacometrics Syst Pharmacol 2017;6:604–13.
24. Bender BC, Schindler E, Friberg LE. Population pharmacokinetic–pharmacodynamic modelling in oncology: a tool for predicting clinical response. Br J Clin Pharmacol 2015;79:56–71.
25. Shirao K, Nishida T, Doi T, Komatsu Y, Muro K, Li Y, et al. Phase I/II study of sunitinib malate in Japanese patients with gastrointestinal stromal tumor after failure of prior treatment with imatinib mesylate. Invest New Drugs 2010;28:866–75.
26. Maki RG, Fletcher JA, Heinrich MC, Morgan JA, George S, Desai J, et al. Results from a continuation trial of SU11248 in patients (pts) with imatinib (IM)-resistant gastrointestinal stromal tumor (GIST). J Clin Oncol 2005;23:9011.
27. Beal SL, Sheiner LB, Boeckmann A, Bauer RJ, editors. NONMEM 7.3.0 Users Guides. Hanover, MD: Icon Development Solutions; 2009.
28. 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.
29. Baron KT. mrgsolve: simulate from ODE-based population PK/PD and systems pharmacology models. Tariffville, CT: Metrum Research Group; 2017.

30. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2009. Available from: https://ggplot2-book.org/.
31. Que Y, Liang Y, Zhao J, Ding Y, Peng R, Guan Y, et al. Treatment-related adverse effects with pazopanib, sorafenib and sunitinib in patients with advanced soft tissue sarcoma: a pooled analysis. Cancer Manag Res 2018;10:2141–50.
32. McDonald JH. Handbook of biological statistics. 3rd ed. Baltimore, Maryland: Sparky House Publishing; 2014.
33. Kowalski KG, Hutmacher MM. Design evaluation for a population pharmacokinetic study using clinical trial simulations: a case study. Stat Med 2001;20:75–91.
34. Sheiner LB, Beal S, Rosenberg B, Marathe VV. Forecasting individual pharmacokinetics. Clin Pharmacol Ther 1979;26:294–305.
35. Keizer RJ, Karlsson MO, Hooker A. Modeling and simulation workbench for NONMEM: tutorial on Pirana, PsN, and Xpose. CPT Pharmacometrics Syst Pharmacol 2013;2:e50.
36. Centanni M, Friberg LE. Model-based biomarker selection for dose individualization of tyrosine-kinase inhibitors. Front Pharmacol 2020;11:316.
37. Khosravan R, Motzer RJ, Fumagalli E, Rini BI. Population pharmacokinetic/pharmacodynamic modeling of sunitinib by dosing schedule in patients with advanced renal cell carcinoma or gastrointestinal stromal tumor. Clin Pharmacokinet 2016;55:1251–69.
38. Kollmannsberger C, Bjarnason G, Burnett P, Creel P, Davis M, Dawson N, et al. Sunitinib in metastatic renal cell carcinoma: recommendations for management of noncardiovascular toxicities. Oncologist 2011;16:543–53.
39. Barrios CH, Herchenhorn D, Chacón M, Cabrera-Galeana P, Sajben P, Zhang K. Safety and efficacy of sunitinib in patients from Latin America: subanalysis of an expanded access trial in metastatic renal cell carcinoma. Onco Targets Ther 2016;9:5839–45.
40. Rini BI, Hutson TE, Figlin RA, Lechuga MJ, Valota O, Serfass L, et al. Sunitinib in patients with metastatic renal cell carcinoma: clinical outcome according to International Metastatic Renal Cell Carcinoma Database Consortium Risk Group. Clin Genitourin Cancer 2018;16:298–304.
41. Jansson-Löfmark R, Hjorth S, Gabrielsson J. Does *in vitro* potency predict clinical efficacious concentrations? Clin Pharmacol Ther 2020 [published online ahead of print, 2020 Apr 10].
42. de Wit D, Guchelaar H-J, den Hartigh J, Gelderblom H, van Erp NP. Individualized dosing of tyrosine kinase inhibitors: are we there yet? Drug Discov Today 2015;20:18–36.
43. Westerdijk K, Krens SD, van der Graaf WTA, Mulder SF, van Herpen CML, Smilde T, et al. The relationship between sunitinib exposure and both efficacy and toxicity in real-world patients with renal cell carcinoma (RCC) and gastrointestinal stromal tumour (GIST). Br J Clin Pharmacol 2020 [published online ahead of print, 2020 May 1].
44. Goulooze SC, Galettis P, Boddy AV, Martin JH. Monte Carlo simulations of the clinical benefits from therapeutic drug monitoring of sunitinib in patients with gastrointestinal stromal tumours. Cancer Chemother Pharmacol 2016;78:209–16.
45. Zietz S, Nicolini C. Mathematical approaches to optimization of cancer chemotherapy. Bull Math Biol 1979;41:305–24.
46. Harrold JM, Parker RS. Clinically relevant cancer chemotherapy dose scheduling via mixed-integer optimization. Comput Chem Eng 2009;33:2042–54.
47. Svensson RJ, Niward K, Davies Forsman L, Bruchfeld J, Paues J, Eliasson E, et al. Individualised dosing algorithm and personalised treatment of high-dose rifampicin for tuberculosis. Br J Clin Pharmacol 2019;85:2341–50.
48. Rini BI, Melichar B, Fishman MN, Oya M, Bair AH, Chen Y, et al. Axitinib dose titration: analyses of exposure, blood pressure and clinical response from a randomized phase II study in metastatic renal cell carcinoma. Ann Oncol 2015;26:1372–7.
49. León-Mateos L, Mosquera J, Aparicio LA. Treatment of sunitinib-induced hypertension in solid tumor by nitric oxide donors. Redox Biol 2015;6:421–5.
50. Sunitinib: Drug information - UpToDate [Internet]. 2017. [cited 2018 Mar 27]. Available from: http://www.kuwaitcancercenter.com/resources/Drug-Monography/SUNitinib.pdf