Citation: CPT Pharmacometrics Syst. Pharmacol. (2017) 6, 449–457; doi:10.1002/psp4.12195

ORIGINAL ARTICLE

Pharmacometric Modeling of Liver Metastases' Diameter, Volume, and Density and Their Relation to Clinical Outcome in Imatinib-Treated Patients With Gastrointestinal Stromal Tumors

E Schindler[1], SM Krishnan[1], RHJ Mathijssen[2], A Ruggiero[3], G Schiavon[2,4] and LE Friberg[1]*

Three-dimensional and density-based tumor metrics have been suggested to better discriminate tumor response to treatment than unidimensional metrics, particularly for tumors exhibiting nonuniform size changes. In the developed pharmacometric modeling framework based on data from 77 imatinib-treated gastrointestinal patients, the time-courses of liver metastases' maximum transaxial diameters, software-calculated actual volumes (V_{actual}) and calculated ellipsoidal volumes were characterized by logistic growth models, in which imatinib induced a linear dose-dependent size reduction. An indirect response model best described the reduction in density. Substantial interindividual variability in the drug effect of all response assessments and additional interlesion variability in the drug effect on density were identified. The predictive ability of longitudinal tumor unidimensional and three-dimensional size and density on overall survival (OS) and progression-free survival (PFS) were compared using parametric time-to-event models. Death hazard increased with increasing V_{actual}. This framework may guide early clinical interventions based on three-dimensional tumor responses to enhance benefits for patients with gastrointestinal stromal tumors (GIST).
CPT Pharmacometrics Syst. Pharmacol. (2017) **6**, 449–457; doi:10.1002/psp4.12195; published online 5 April 2017.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Conventional unidimensional size measurements may not adequately detect treatment response for tumors with nonspherical shapes or nonuniform size changes, such as liver metastases from imatinib-treated GIST. Instead, 3D measurements and structural density assessments may better reflect treatment effect.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ Pharmacometric models characterized the unidimensional, 3D, and density responses of imatinib-treated GIST liver metastases, and compared the predictive ability of each metrics on OS and PFS.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

☑ Three-dimensional may be more sensitive than unidimensional measurements at detecting tumor response to imatinib. Three-dimensional changes predicted OS slightly better than unidimensional changes, whereas density changes had no predictive value.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

☑ These results improve our understanding of the relationships among imatinib exposure, conventional and alternative tumor measurements, and long-term outcomes. They may guide early clinical decisions for patients with GIST, and encourage the incorporation of 3D measurements in prospective clinical trials to test their ability and value vs. unidimensional measurements in predicting outcomes.

Efficacy assessment in cancer clinical trials and clinical practice to evaluate tumor shrinkage or *vice versa* development of disease progression has been standardized through the implementation of criteria for morphological imaging measurements. Identification of responders and nonresponders early after therapy initiation is crucial to trigger treatment modification whenever needed. Highly sensitive methods for tumor response quantification that correlate with clinical outcome are, therefore, required. The most recent and widely-adopted Response Evaluation Criteria in Solid Tumors (RECIST) are based on unidimensional (1D) changes in the sum of longest diameters of target tumor lesions, assessed by magnetic resonance imaging or computed tomography.[1,2] One question raised by the RECIST Working Group when updating the guidelines was whether and when functional or volumetric assessment could supplement or potentially replace RECIST.[2,3]

Choi *et al.*[4,5] proposed novel criteria combining tumor size and density changes that reflect morphological and

[1]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden; [2]Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands; [3]Department of Radiology, Papworth Hospital NHS Foundation Trust, Cambridge University Health Partners, Cambridge, CB23 3RE, United Kingdom; [4]Current address: Translational Science, IMED Oncology, AstraZeneca, Cambridge, UK. *Correspondence: L Friberg (lena.friberg@farmbio.uu.se)
Received 17 January 2017; accepted 22 March 2017; published online on 5 April 2017. doi:10.1002/psp4.12195

structural changes, respectively, to assess response of gastrointestinal stromal tumors (GIST) to imatinib, a multitargeted tyrosine kinase inhibitor.[4,5] In imatinib-treated GIST, the Choi criteria correlated with time-to-tumor progression and disease-specific survival.[6] Contrasting results have been reported for different cancer types regarding the predictive value of Choi criteria on clinical outcome, such as overall survival (OS) or progression-free survival (PFS).[7–12]

As tumors are three-dimensional (3D) structures, using a single dimension might not accurately reflect tumor burden and may underestimate shrinkage or overestimate growth, especially for lesions that are nonspherical and/or display nonuniform size changes during treatment, such as primary tumors[13] and liver metastasis[14,15] from GIST or in tumors whose main changes occur in a plan other than axial. At the time of the RECIST update in 2009 (RECIST 1.1), the lack of rigorous clinical validation and of widespread access to these newer imaging approaches prevented their implementation and adoption. However, in recent years, a number of publications have demonstrated the high reproducibility and interobserver reliability of volume measurements using semiautomated tumor segmentation technique implemented in commercially available[15–18] or in-house software.[19] The estimated ellipsoidal volume has been suggested as an approximation for software-calculated actual volume to overcome limitations related to software availability and has shown good correlation to actual volume.[15,18] Moreover, for a number of cancers 3D assessments enable to capture finer changes in tumor size,[17,18] and better correlates to long-term clinical outcome, such as OS,[2,18,21] compared to 1D measurements. Such results were supported by Schiavon *et al*.,[15] showing that ellipsoidal volume provides a satisfactory approximation for the actual volume of liver metastases in imatinib-treated GIST. The 3D changes could be detected more frequently than 1D changes after 3 months of treatment. Moreover, volume-based criteria tended to be better associated to OS than RECIST. Noteworthy, longitudinal analysis of tumor volume on a continuous scale are only rarely reported[21]; most published results are based on landmark analyses, in which the correlation between the often-categorized change in tumor burden after a fixed treatment duration and clinical outcome is investigated, resulting in a loss of information.

Pharmacometric models have proved their value in leveraging data collected during clinical trials and clinical practice.[22–24] They allow for characterization of the tumor response time-course as a function of drug exposure, quantification of interindividual variability (IIV), interlesional variability (ILV),[25] and residual variability in the response, and investigation of the relationships between longitudinal tumor size and time-to-event outcome. In the present study, we developed a pharmacometric modeling framework to (i) characterize the time-course and quantify the IIV and ILV of tumor 1D, 3D, and density response of liver metastases in imatinib-treated patients with GIST and (ii) investigate the predictive ability of these size and density metrics on OS and PFS.

METHODS

Study population and data

The tumor data were available from 77 patients with GIST treated with first-line imatinib and involved in two retrospective, noninterventional studies based on historical data published previously.[15,26] Imatinib was administered orally at a starting dose of 400 mg ($N = 74$) or 800 mg ($N = 3$) daily until unacceptable toxicity or disease progression (assessed as progressive disease by RECIST). Any dose escalation or reduction was based on safety and efficacy and recorded at routine clinical assessments. Information on subsequent systemic therapy (e.g., sunitinib and regorafenib) was also available. One or two liver metastases (target lesions) per patient were retrospectively selected to be followed on computed tomography scans performed from initiation of imatinib treatment (baseline) and at least once during treatment (at ~3, 6, and 12 months), with a median time for tumor follow-up of 360 days (range, 82–495 days). Tumor 1D measurements of target lesion(s) consisted of the maximum transaxial diameter (MTD), as per RECIST version 1.1 definition. Tumor 3D measurements of the same lesions included the segmented volume (V_{actual}), calculated using the semi-automated segmentation by "syngo computed tomography Oncology" software version 2009E (Siemens Medical Solutions), and the calculated ellipsoidal volume ($V_{ellipsoid} = \frac{\pi}{6} \cdot MTD \cdot d_2 \cdot d_3$, where MTD, d_2, and d_3 are three orthogonal diameters).[26] For plotting purposes, the spherical volume ($V_{spherical} = \frac{\pi}{6} \cdot MTD^3$) was calculated assuming that lesions are perfect spheres. Tumor density (in Hounsfield units [HUs]) was recorded for each lesion by measuring the computed tomography attenuation coefficient (see Schiavon *et al*.[26] for details on the computation of these measurements). OS and PFS data were collected for a median of 4.5 years (range, 0.79–13 years) and 3.4 years (range, 0.25–13 years), respectively, after treatment initiation. PFS was defined as the time from first imatinib dose to discontinuation of treatment for any reason, including disease progression (assessed by RECIST), drug-related toxicities, or death. Censoring was defined as loss to follow-up or nonoccurrence of death or progression at the end of the follow-up period, for OS and PFS, respectively. In the analysis presented here, OS and PFS data collection was updated and closed on February 17, 2015. **Table 1** provides a summary of patients' characteristics and available data. Ethical standards were followed and approvals were obtained as described earlier.[26]

Tumor models

MTD, V_{actual}, ellipsoidal volume ($V_{ellipsoid}$), and density models were developed separately before being combined into a joint model to explore correlations and to simplify the model structure where appropriate. Exponential IIV and ILV were explored in all non-negative model parameters (Eq. 1).[25]

$$\theta_i = \begin{cases} \theta \cdot \exp(\eta_i + \kappa_{i1}) & \textit{if lesion } 1 \\ \theta \cdot \exp(\eta_i + \kappa_{i2}) & \textit{if lesion } 2 \end{cases} \quad (1)$$

θ_i represents the parameter for the i^{th} individual and θ the typical parameter value in the population. η_i is the term describing IIV, common to all lesions but specific to each individual ($\eta_i \sim N(0, \omega^2)$). κ_{i1} and κ_{i2} are the ILV terms for lesion 1 and lesion 2, respectively, assumed to be normally distributed with mean 0 and a common variance π^2.

Table 1 Summary of patients' characteristics and available data

No. of patients, total/with 2 lesions	77/60
Gender, male/female, no. (%)	47/30 (61/39)
Age at start of imatinib, years, median (range)	62 (34–83)
Imatinib starting dose, mg (n)	400/800 (74/3)
Second line of therapy and beyond (yes/no), no. (%)	35/42 (45/55)
Patients with dose escalation[a], no. (%)	30 (39)
Escalated dose, mg, median (range)	800 (600–1200)
Patients with dose reductions[a], no. (%)	8 (10)
Reduced dose, mg, median (range)	300 (200–300)
Events for overall survival (censored/death), no. (%)	34/43 (45/55)
Events for progression-free survival (censored/progression), no. (%)	27/50 (35/65)

[a]Taking into account first dose alteration only.

Residual unexplained variability was described by additive models on a logarithmic scale.

Maximum transaxial diameter, actual volumes, and ellipsoidal volume models

Tumor models in which imatinib-induced tumor size reductions were developed describe log-transformed MTD, V_{actual}, and $V_{ellipsoid}$ data. Various tumor growth models were explored, including exponential, zero-order, Gompertz, and logistic functions.[27,28] Linear, power, and maximum effect functions of imatinib daily dose were compared to describe the dose-response relationship. A time-dependent mono-exponential decay in drug effect was evaluated to explain tumor regrowth following initial shrinkage. Addition of an effect compartment was also tested. Semiparametric distributions[29] and mixture models were investigated to describe the observed bimodal distribution in baseline MTD, V_{actual}, and $V_{ellipsoid}$.

Tumor density model

Indirect response models[30] in which imatinib inhibits the input or stimulates the output of density response were investigated to describe log-transformed tumor density data, which tended to decrease during therapy. Linear and nonlinear (power, maximum effect) drug effect functions driven by imatinib daily dose were tested. A potential delay in affecting density changes was evaluated through the addition of an effect compartment. A potential drug-independent increase in density over time was assessed using time-dependent linear or nonlinear (that plateaus over time, as apparent from some observed lesion profiles) disease progression functions applied on the production rate.

Overall survival and progression-free survival models

Parametric time-to-event (TTE) models were developed to investigate the effect of potential predictors of OS and PFS hazards ($h(t)$). The baseline hazard ($h_0(t)$) was explored using exponential, Weibull, Gompertz, log-normal, and log-logistic models.[31] A sequential analysis similar to the population pharmacokinetic parameters and data[32] approach was initially considered but led to model instability, preventing likelihood ratio testing. Therefore, an approach similar to the individual pharmacokinetic parameter approach was used,[32] in which individual empirical Bayes estimates from the final joint tumor model were used to predict individual tumor size or density time-courses, which were then investigated as predictors in the TTE models. Baseline predictors included patient characteristics (age and gender) and model-predicted (log-transformed) baseline MTD, V_{actual}, $V_{ellipsoid}$, and tumor density. Time-varying predictors tested on OS consisted of (log-transformed) model-predicted tumor time-courses ($S(t)$), and absolute ($dS(t)$) and relative ($S_{rel}(t)$) change from baseline over time, where S is MTD, V_{actual}, or $V_{ellipsoid}$. Additionally, relative change in size and density up until 1.5, 3, 6, and 12 months for OS and 1.5 and 3 months for PFS were evaluated. As the PFS definition is partly based on MTD changes, only early size-related predictors (1.5 or 3 months) were considered in the PFS analysis to minimize potential confounding between the predictor and the outcome. When patients presented with two lesions, MTD, V_{actual}, and $V_{ellipsoid}$ were summed across lesions, whereas density was averaged over lesions. Predictors were evaluated alone (univariate analysis) and in combination (multivariate analysis). Predictors related to different size metrics were not tested in combination, but size-related predictors could be combined with density-related predictors. For each endpoint, a competing TTE model described censoring.

Model development and evaluation

Models were developed using the nonlinear mixed-effects modeling software (NONMEM version 7.3).[33] The first-order conditional estimation method with interaction was used for parameter estimation. R (version 3.1) was used for data management and graphical analysis. Model development and evaluation was aided by Perl-speaks-NONMEM (PsN) toolkit (version 4.5), the R-based package Xpose (version 4), and Pirana (version 2.9.2).[34]

Model performance was evaluated based on objective function value (OFV; $-2\cdot$log-likelihood) and graphical diagnostics. Upon addition of one parameter (1 degree of freedom (df)), an OFV decrease of at least 6.63 ($P < 0.01$) for tumor models (repeated measures within subjects), and an OFV of at least 3.84 ($P < 0.05$) for TTE models (sparser data – single event per individual), was considered as statistically significant. Relative standard errors (RSEs) of parameter estimates were obtained from the NONMEM Sandwich matrix for continuous data and from the R matrix for TTE models. The predictive performance of continuous models was evaluated by visual predictive checks (VPCs), in which 95% confidence intervals (CIs) of the 10th, 50th, and 90th percentiles from 1,000 simulated datasets were compared to the corresponding percentiles from the observed data. TTE models' predictive performance was assessed by Kaplan–Meier VPCs, in which the 95% confidence interval based on 200 simulated datasets was compared to the observed Kaplan–Meier curve.

RESULTS

Data

Tumor data consisted of 502 observations for each size metrics and 496 density observations, collected from 136 liver lesions in 77 patients (maximum 2 per patient). Lesions were numbered so that lesion 1 has the largest MTD at baseline.

Table 2 Parameter estimates and their uncertainty in the final tumor models for maximum transaxial diameter, actual volumes, and ellipsoidal volume

		MTD		V_{actual}		$V_{ellipsoid}$	
Parameter (unit)		**Typical value (RSE%)**	**IIV, CV% (RSE%)**	**Typical value (RSE%)**	**IIV, CV% (RSE%)**	**Typical value (RSE%)**	**IIV, CV% (RSE%)**
$S_{0,\ pop\ 1}$, mm or mL[a]	Lesion 1	76.6 (12)	47 (11)[b]	161 (34)	140 (12)[b]	187 (35)	140 (12)[b]
	Lesion 2	41.9 (15)	47 (11)[b]	29.7 (46)	140 (12)[b]	33.4 (46)	140 (12)[b]
$S_{0,\ pop\ 2}$, mm or mL[a]	Lesion 1	20.9 (6.3)	29 (13)[c]	3.45 (16)	76 (13)[c]	3.93 (17)	78 (12)[c]
	Lesion 2	14.2 (6.7)	29 (13)[b]	1.21 (18)	76 (13)[b]	1.27 (19)	78 (12)[b]
S_{max}, mm or mL[a]	Lesion 1	171 (8.4)	-	1190 (43)	-	1230 (57)[d]	-
	Lesion 2	125 (3.7)	-	540 (12)	-	588 (18)	-
K_G (week^{-1})		0.00176 (47)	170 (18)	0.00861 (38)	135 (20)	0.00882 (43)	140 (27)
λ (week^{-1})		0.0475 (35)	-	0.0469 (28)	-	0.0508 (37)	-
$K_{drug,S}$ (week^{-1})		0.0124 (36)	77 (20)	0.0547 (21)	56 (29)	0.0610 (24)	42 (36)
$P_{pop\ 1}$		0.348 (18)[c]	-	0.348 (18)[c]	-	0.348 (18)[c]	-
RUV (%)		14.0 (6.6)	-	36.8 (8.0)	-	43.3 (7.5)	-

CV, coefficient of variation; λ, resistance development rate constant; IIV, interindividual variability; $K_{drug,S}$, slope of the linear drug effect; K_G, tumor growth rate constant; MTD, mean transaxial diameter; $P_{pop\ 1}$, probability of belonging to mixture subpopulation 1; RSE, relative standard error; RUV, residual unexplained variability S_{max}, carrying capacity; $S_{0,\ pop\ n}$, baseline size in mixture subpopulation n; V_{actual}, actual volume; $V_{ellipsoid}$, ellipsoidal volume.
Lesions were numbered such that lesion 1 corresponds to the one with the largest MTD at baseline.
[a]Millimeters for MTD, mL for V_{actual} and $V_{ellipsoid}$.
[b]IIV term variance shared between lesions, within each subpopulation and model.
[c]Common parameter to MTD, V_{actual}, and $V_{ellipsoid}$ models.
[d]Ninety-five percent confidence intervals obtained from log-likelihood profiling was (685;3260).

Joint tumor model for maximum transaxial diameter, actual volumes, ellipsoidal volume, and density

A tumor model with a logistic growth function in which imatinib induces tumor size reduction (Eq. 2) best characterized the individual lesion data from all three tumor sized metrics. Logistic growth fitted the data better than exponential growth (objective function value difference [dOFV]; −10.3, −13.4, and −10.2 for MTD, V_{actual}, and $V_{ellipsoid}$, respectively). A linear effect with slope $K_{drug,S}$ driven by imatinib daily dose, normalized by the median dose (400 mg), best described the dose-response relationship. Moreover, a time-dependent mono-exponential function with rate constant λ described a drug effect washout. Addition of an effect compartment did not improve model fit.

$$\frac{dS}{dt} = K_G \cdot S(t) \cdot \left(1 - \frac{S(t)}{S_{max}}\right) - K_{drug,S} \cdot \frac{Dose}{400} \cdot e^{-\lambda \cdot t} \cdot S(t) \qquad (2)$$

S(t) represents the lesion size (MTD, V_{actual}, or $V_{ellipsoid}$) at time t. K_G is the tumor growth rate constant and S_{max} the maximal size above which the lesion cannot grow (i.e., carrying capacity). A mixture model with two subpopulations of patients described the bimodal distribution of baseline lesion size (S_0), with subpopulation 1 (estimated probability of 0.348) having lesions with larger baseline V_{actual}, $V_{ellipsoid}$, and MTD than subpopulation 2. Of note, 22% of patients with one target lesion were assigned to subpopulation 1 vs. 39% for patients with two target lesions. For a typical patient with two lesions, the estimated total S_0 was 3.4, 41, and 42 times larger in subpopulation 1 than in subpopulation 2, for MTD, V_{actual}, and $V_{ellipsoid}$, respectively. For each mixture and size metrics, S_0 was associated with an ILV term that was specific to each lesion but shared a common variance for the two lesions. The typical total S_{max} estimate for an individual with two lesions was 296 mm for MTD, 1,730 mL for V_{actual}, and 1,818 mL for $V_{ellipsoid}$. For V_{actual} and $V_{ellipsoid}$, an alternative parameterization for S_{max} was considered, in which the carrying capacity for the total volume of lesion 1 and 2 ($S_{max,tot}$) was estimated together with the fraction of $S_{max,tot}$ for lesion 1; however, this resulted in model instability and large parameter uncertainty and was not retained. All V_{actual} and $V_{ellipsoid}$ estimates were similar, with $V_{ellipsoid}$ estimates within 100–116% of V_{actual} estimates (**Table 2**). $K_{drug,S}$ estimates were fourfold to fivefold larger in V_{actual} and $V_{ellipsoid}$ models compared to the MTD model; IIV but no ILV was identified in K_G and $K_{drug,S}$.

An indirect response model in which imatinib stimulates the output of response through a linear drug effect with slope $K_{drug,D}$ driven by imatinib daily dose, normalized by the median dose (400 mg), best described tumor density time-course ($D(t)$) (Eq. 3). Addition of an effect compartment did not improve model fit (dOFV = −1.14). Accounting for disease progression using an increase in the production rate that levels off over time statistically significantly improved the OFV (dOFV = −13.3); however, because of that the OFV decrease was highly dependent on two individuals, model instability and negligible improvement in VPCs, disease progression was not retained. A Box-Cox transformation with an estimated shape parameter of −1.06 was applied to baseline density (D_0) sum of η and κ to account for the skewed random effects distribution. The IIV magnitude of $K_{drug,D}$ was larger than the ILV magnitude (120 vs. 53% CV). No IIV or ILV was identified in the rate constant for loss of response (k_{out}).

$$\frac{dD}{dt} = R_{in} - k_{out} \cdot \left(1 + K_{drug,D} \cdot \frac{Dose}{400}\right) \cdot D(t) \qquad (3)$$

The rate constant for production of response was parameterized as $R_{in} = k_{out} \cdot D_0$. The derived mean turnover time ($MRT = 1/k_{out}$) was 75 days. Estimated correlations

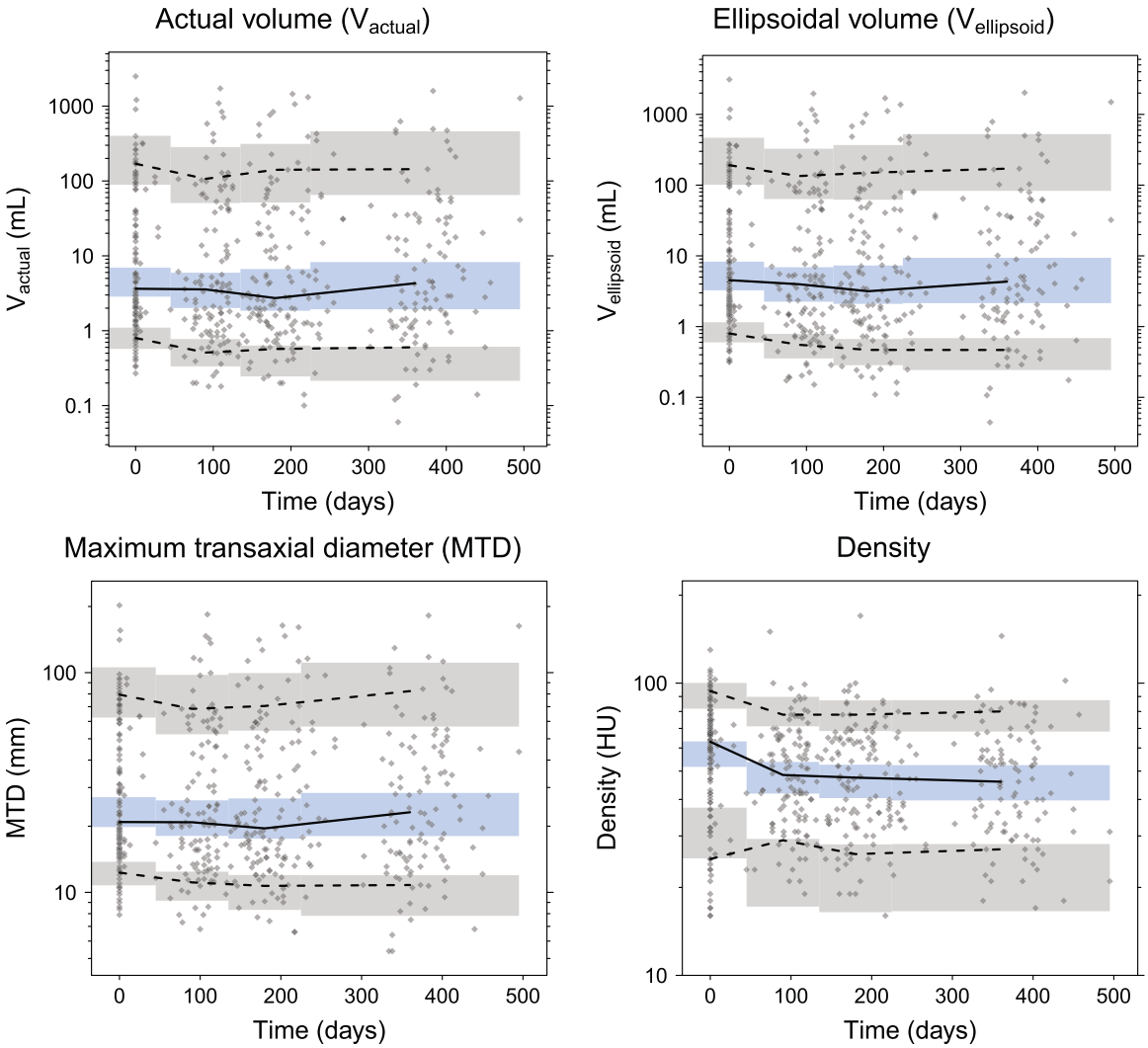

Figure 1 Visual predictive checks of the joint tumor model for actual volume (V_{actual}), ellipsoidal volume ($V_{ellipsoid}$), maximum transaxial diameter (MTD), and density models. Median (solid line), 10th and 90th percentiles (dashed lines) of the observed data are compared to the 95% confidence intervals (shaded areas) for the median (blue), 10th and 90th percentiles (gray) of the simulated data, based on 1,000 simulations.

between density model parameters and MTD, V_{actual}, or $V_{ellipsoid}$ model parameters were <20%.

VPCs indicated appropriate predictive performance of the joint model (**Figure 1 and Supplementary Figures S1–S4**). **Tables 2 and 3** present the joint tumor model parameter estimates and their uncertainty. Parameter uncertainty was within acceptable range (≤47% RSE), except for $S_{max,lesion\ 1}$ in the $V_{ellipsoid}$ model (56% RSE), for which 95% confidence interval from log-likelihood profiling[35] was 685–3,260 mL.

Overall survival and progression-free survival models

A lognormal distribution with estimated mean μ_{OS} and μ_{PFS} and SD σ_{OS} and σ_{PFS}, for OS and PFS respectively, best described the baseline hazard ($h_0(t)$).

In the univariate OS analysis, the most statistically significant predictor was log-transformed $V_{actual}(t)$ (dOFV = −8.09). $V_{actual}(t)$ (dOFV = −7.09) and $dV_{actual}(t)$ (dOFV = −4.26) were also significant but to a lower degree. Other statistically significant predictors identified in the univariate analysis are summarized in **Supplementary Material Table S1**. None of the density-related variables or patient demographics predicted OS. In the multivariate OS analysis, no other predictor improved the fit further when log-transformed $V_{actual}(t)$ had been added. This model was concluded in the final OS model (Eq. 4).

$$h(t) = \frac{\frac{1}{t \cdot \sigma_{OS} \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{\log(t) - \mu_{OS}}{\sigma_{OS}}\right)^2}}{1 - \Phi\left(\frac{\log(t) - \mu_{OS}}{\sigma_{OS}}\right)} \cdot e^{\theta_{V_{actual}} \cdot \log(V_{actual}(t))} \quad (4)$$

$\Phi(\cdot)$ is the standard normal cumulative distribution function. The coefficient for the effect of $\log(V_{actual}(t))$ on the hazard, $\theta_{V_{actual}}$, was estimated to 0.190, denoting an increased death hazard with increasing V_{actual}. **Figure 2** (left panel) illustrates the hazard ratio (HR) over the range of observed V_{actual}, using the typical total S_0 (2 lesions, subpopulation 1) as reference. For instance, the HR is estimated to 1.52

Table 3 Parameter estimates and their uncertainty in the final tumor density model.

Parameter (unit)	Typical value (RSE%)	IIV – CV% (RSE%)	ILV – CV% (RSE%)
D_0 (HU)	59.0 (5.7)	30 (10)	18 (19)
Box-Cox D_0	−1.06 (47)[a]	-	-
k_{out} (week^{-1})	0.0935 (32)	-	-
$K_{drug,D}$	0.154 (29)	120 (17)	53 (38)
RUV (%)	20.6 (6.8)	-	

CV, coefficient of variation; D_0, baseline tumor density; HU, Hounsfield unit; IIV, interindividual variability; ILV, interlesion variability; K_{drug}, density reduction constant; k_{out}, rate constant for loss of tumor density; RSE, relative standard error; RUV, residual variability.
[a]The 95% confidence intervals obtained from log-likelihood profiling was (-2.01;-0.397) for Box-Cox D_0.

at the typical total V_{actual} S_{max} (1,730 mL) in subpopulation 1, and to 0.494 at the typical total V_{actual} S_0 value (4.66 mL) in subpopulation 2.

Noteworthy, in the OS multivariate analysis, a lower OFV was achieved by combining log-transformed baseline and relative change from baseline up to 3 months in V_{actual} (dOFV = −13.7) or $V_{ellipsoid}$ (dOFV = −13.3), whereas for MTD the corresponding combination resulted in a smaller improvement in the model fit (dOFV = −9.31); however, these multivariate models requested one extra model parameter compared to the final model.

In the univariate PFS analysis, the most statistically significant predictor was V_{actual} relative change from baseline up to 3 months ($V_{actual,rel,3m}$, dOFV = −13.9). All other investigated size-related predictors were also significant, with V_{actual} and $V_{ellipsoid}$ predictors always being better than corresponding MTD predictors. None of the density-related variables could predict PFS. The model fit improved further when log-transformed V_{actual} baseline ($\log[V_{actual,0}]$, dOFV = −12.0) was added (Eq. 5).

$$h(t)=\frac{\frac{1}{t\cdot\sigma_{PFS}\cdot\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\cdot\left(\frac{\log(t)-\mu_{PFS}}{\sigma_{PFS}}\right)^2}}{1-\Phi\left(\frac{\log(t)-\mu_{PFS}}{\sigma_{PFS}}\right)}\cdot e^{\theta_{V_{actual,rel,3m}}\cdot V_{actual,rel,3m}+\theta_{V_{actual,0}}\cdot log\left(V_{actual,0}\right)} \quad (5)$$

The estimated coefficients for $V_{actual,rel,3m}$ ($\theta_{V_{actual,rel,3m}}$) and $\log(V_{actual,0})$ ($\theta_{V_{actual,0}}$) effects indicate that hazard increases with increasing V_{actual} and larger baseline (**Table 4**). The estimated HR was 0.920 (95% confidence interval = 0.892–0.948) for every 10% decrease in V_{actual}.

Kaplan–Meier VPCs (**Figure 3**) show satisfactory predictive performance of the OS and PFS models, even though the PFS model tended to underpredict the survival rate toward the end of the study period. Parameter estimates and their uncertainty for both models are reported in **Table 4**.

DISCUSSION

In this pharmacometric modeling analysis of patients with GIST, we characterized the time-course of liver metastases response to first-line imatinib treatment, including three size metrics – the unidimensional MTD, the software-calculated volume V_{actual}, and the estimated ellipsoidal volume $V_{ellipsoid}$ – as well as tumor density, which informs on the tumor structural status. Log-transformed V_{actual} time-course best predicted OS. Interestingly, $\log(V_{actual}(t))$ provided a better fit than $V_{actual}(t)$, denoting a nonlinear relationship between the HR and V_{actual} (**Figure 2** (left panel)). The V_{actual} relative change from baseline up to 3 months together with log-transformed V_{actual} baseline best predicted PFS.

The same structural tumor model could be applied to all three sized metrics. The binomial distribution of the observed baseline was accounted for by a mixture model, in which the probability of belonging to a subpopulation with typically larger tumors than in the other subpopulation was estimated to 34.8%. The assignment to one or the other subpopulation was not related to any of the available demographic variables (age, gender, or clinical center). A logistic growth model, in which the tumor growth rate decreases as the tumor size approaches the carrying capacity, described the data better than an exponential growth (as expected, presumably due to the dependency of the tumor growth from a limited vascular supply). The typical total carrying capacity (both lesions accounted for) was estimated to 1,730 and 1,818 mL for V_{actual} and $V_{ellipsoid}$, respectively; these values are similar to literature values of liver volume in living donors (*c.a.* 1,500 mL)[36] and to the baseline liver volume in a similar GIST patient population[14] (median = 1,755; range = 1,112–3,354 mL), of which some of the patients reported in this study were part, and, therefore, clinically plausible. When tumor size is very small compared to the carrying capacity, logistic growth can be approximated by an exponential growth with rate constant K_G. The K_G estimate for MTD corresponds to a typical doubling time of small tumors of 7.4 years, vs. 1.5 years for V_{actual} and $V_{ellipsoid}$. The latter is comparable with the value of 1.0 year reported for untreated GIST before surgical resection.[37] The slope of the drug effect was estimated to be lower for MTD than for V_{actual} and $V_{ellipsoid}$; 3D measurements may, therefore, be more sensitive to detect imatinib-induced changes compared with 1D measurements. This is illustrated in **Figure 2** (middle panel), in which the predicted typical ratio to baseline decreases less for MTD than for 3D measurements. Large IIV (range, 165–

Table 4 Parameter estimates and their uncertainty in the final overall survival and progression-free survival models.

OS model		PFS model	
Parameter	**Estimate (% RSE)**	**Parameter**	**Estimate (% RSE)**
μ_{OS}	1.42 (15)	μ_{PFS}	1.19 (12)
σ_{OS}	8.46 (5.0)	σ_{PFS}	7.79 (3.4)
$\theta_{Vactual}$	0.190 (36)	$\theta_{Vactual,rel,3m}$	0.836 (19)
		$\theta_{Vactual,0}$	0.239 (28)

OS, overall survival; PFS, progression-free survival; RSE, relative standard error; μ_{OS}, μ_{PFS}, mean of the log-normal hazard model for OS and PFS, respectively; σ_{OS}, σ_{PFS}, SD of the log-normal hazard model for OS and PFS; $\theta_{Vactual}$, coefficient of the effect of log-transformed actual volume time-course on the hazard of OS; $\theta_{Vactual,rel,3m}$, coefficient of the effect of actual volume relative change from baseline up to 3 month on the hazard of PFS; $\theta_{Vactual,0}$, coefficient of the effect of log-transformed actual volume baseline on the hazard of PFS.

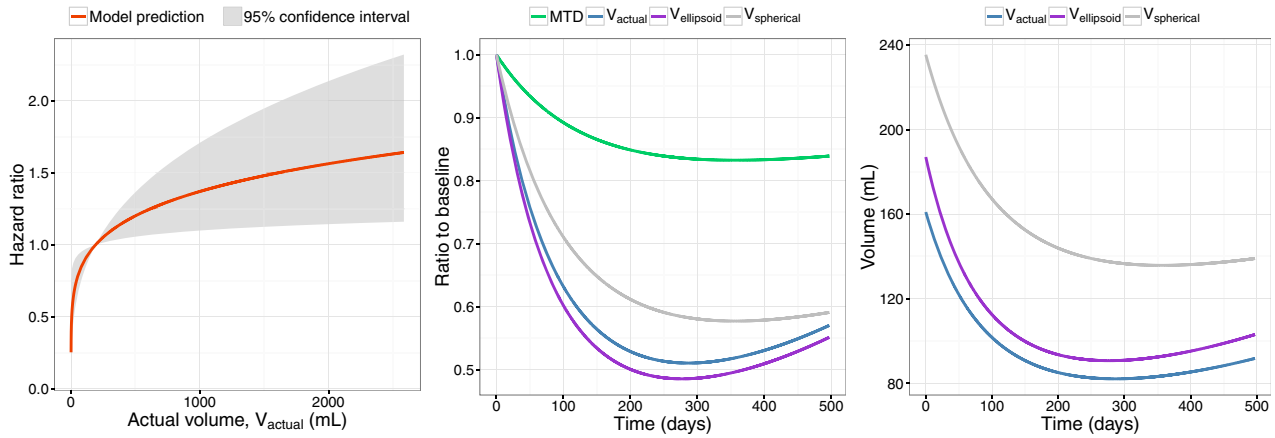

Figure 2 Model predictions of the hazard ratio (HR) from the final overall survival (OS) model vs. the actual volume (V_{actual}) (left), and of the ratio to baseline (middle) and volumes (right) vs. time from the final joint tumor model. The HR (left) is plotted for the range of observed V_{actual} (0.14–2586 mL), taking the typical baseline V_{actual} in the mixture subpopulation with larger lesions at baseline (190.7 mL) as a reference. The gray shaded area shows the 95% confidence interval based on parameter uncertainty in the final OS model. Model-predicted ratio to baseline (middle) and volumes (right) vs. time are plotted for the largest lesion in a typical individual belonging to the mixture subpopulation with larger lesions at baseline and treated with 400 mg imatinib daily. The spherical volume ($V_{spherical}$) time-course was derived from the maximum transaxial diameter (MTD) time-course ($V_{spherical} = \pi/6;\cdot MTD/10^3$) assuming that lesions are perfect spheres. $V_{ellipsoid}$: ellipsoidal volume.

170% CV) was quantified on the growth rate constant of all size metrics, and moderate to large IIV on the slope of the imatinib dose effect (42–120% CV). Part of the latter variability may have been explained by differences in pharmacokinetics, if available. ILV was not found to be statistically significant in any of the model parameters, except baseline. Accounting for ILV would enable to test individual lesion response on survival, and could result in more accurate predictions of tumor dynamics, which may be of particular importance when extrapolating the prediction beyond the tumor assessment period.[25]

Tumor density time-course was described by an indirect response model with stimulation of the response output. Tumor density decrease reflects structural changes secondary to treatment (such as inflammation, tumor necrosis, and reduced vascularization) that can induce paradoxical size increase, which may precede size reduction. The weak correlations between density and size-related parameters indicate that structural changes may occur independently of size changes and that both measurements may provide distinct information on tumor response. Substantial ILV (53% CV) was identified on the dose effect parameter in the density model, indicating differences in tumor sensitivity to imatinib and inter-tumor heterogeneity (e.g., due to secondary mutations or different degree of necrosis post-treatment impacting tumor density).[38]

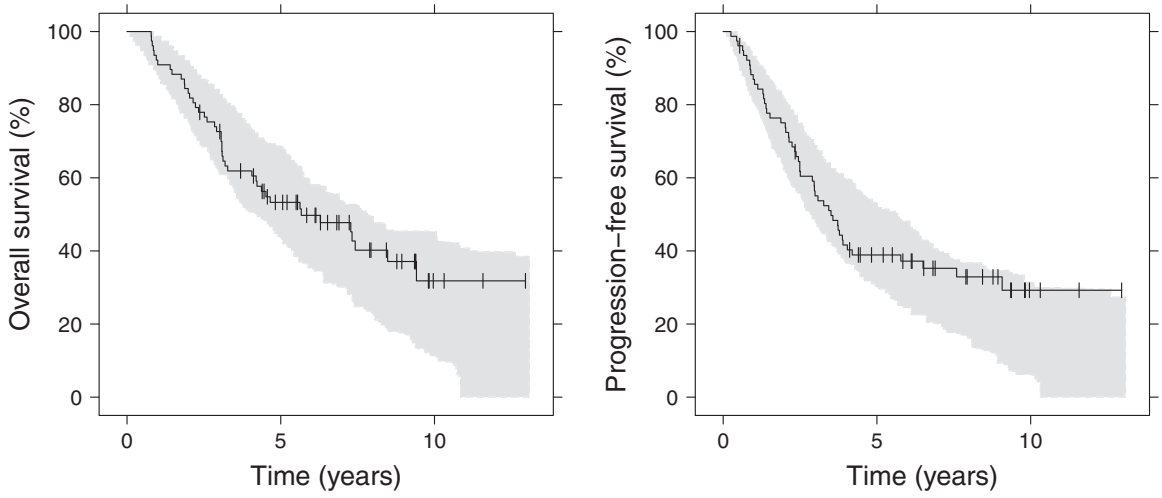

Figure 3 Kaplan–Meier visual predictive checks for the final overall survival (left) and progression-free survival (right) models. The observed Kaplan–Meier curve (black line) is compared to the 95% confidence interval (shaded area) derived from model simulations (200 samples). Vertical black lines represent censored observations.

In the joint tumor model, exploratory analysis showed, as expected, that large correlations (up to 99%) existed between individual model parameters in the MTD, V_{actual}, and $V_{ellipsoid}$ models. However, as the aim of our analysis was to evaluate the predictive ability of each size metrics for OS and PFS independently, no further model simplification was taken in order to retain flexibility in tumor size predictions.

Semi-automated measurements of V_{actual} require software for postprocessing of images, and staff to verify and possibly adjust tumor delineation. In clinical settings, this may be impractical due to software cost and restricted availability, and lack of user experience and time resources. In the future, with the introduction in the clinical arena of deep learning-based software, this challenge can be easily overcome. For well-defined tumors (whose borders can be easily identified), $V_{ellipsoid}$ may be easily and more efficiently obtained than V_{actual} and offer an alternative to account for nonspherical shapes or nonuniform changes in size. The small differences between V_{actual} and $V_{ellipsoid}$ model parameter estimates are consistent with earlier findings in which $V_{ellipsoid}$ was highly correlated to V_{actual} and only overestimated by 9% on average.15 **Figure 2** (right panel) further illustrates the similarity between model-predicted V_{actual} and $V_{ellipsoid}$ time-courses. Moreover, consistent with previous findings, this figure highlights the fact that assuming a spherical volume ($V_{spherical}$) calculated based on MTD would tend to overestimate the actual volume. **Figure 2** (middle panel) also shows that the models predict that the typical tumor size nadir occurs ~80 days earlier for the volume measurements compared to MTD-based tumor sizes, which could be clinically relevant.

To the authors' knowledge, this is the first time that the predictive ability of longitudinal tumor 1D and 3D size and density on clinical outcome has been compared. Despite the differences in goodness-of-fit criteria or predictive ability were marginal, the 3D measurements resulted in a better fit to the OS and PFS data than the 1D measurements. This is in line with the results from the standard statistical analysis by Schiavon *et al.*,[26] which showed that 3D criteria (based on a categorization of tumor response) seemed to overall better correlate to OS than 1D criteria when assessed after ~6 or 12 months of treatment. Contrasting results have been reported regarding the predictive ability of Choi's criteria, a combination of size and density responses, on clinical outcome.[5,8–12] Here, none of the density-related metrics could predict OS or PFS when tested alone or in addition to size-related predictors. An alternative to the final OS model, which includes $\log(V_{actual}(t))$ as predictor, is to use a combination of V_{actual} baseline and its relative change from baseline over time up to 3 months. This could help predicting long-term clinical outcome (in this study, up to 13 years of treatment) using early tumor assessments; however, these results should be interpreted with caution as the study population (77 patients) was small and OS modeling did not account for subsequent therapies and tumor heterogeneity, which may potentially confound the estimated changes in size. Finally, a maximum of two liver metastases were available in this analysis and clinical outcome may be affected by other (or new) tumor lesions in the liver or other organs (as per RECIST version 1.1).

In summary, the developed models adequately described the longitudinal 1D, 3D, and density data of liver metastases in imatinib-treated patients with GIST. All tumor responses were associated with large IIV. The ILV was identified in imatinib effect on density but not size metrics. This analysis provides an insight into the value of 3D and density measurements in response evaluation of GIST to imatinib therapy. The 3D measurements showed promising predictive value of OS, although the results should be validated in a larger independent study population. The present results encourage the incorporation of 3D measurements in prospective clinical trials to evaluate their ability and value vs. unidimensional measurements (RECIST) in predicting clinical outcomes. The developed volume and density models may be used to leverage data from other cancer types and drug therapies, especially where changes in size are believed to be nonuniform. The developed modeling framework provides better understanding of the relationship among drug exposure, short-term tumor response, and long-term clinical outcome, and may guide early clinical decisions and interventions to enhance benefits for patients with GIST.

Acknowledgments. This work was supported by the Swedish Cancer Society.

Conflict of Interest. As Deputy Editor-in-Chief of *CPT: Pharmacometrics & Systems Pharmacology*, L.E.F. was not involved in the review or decision process for this article. G.S. is employee of AstraZeneca. E.S., S.M.K., R.H.J.M., and A.R. declared no conflict of interest.

Author Contributions. Emilie Schindler and Sreenath M. Krishnan share first authorship. E.S., S.M.K., R.H.J.M., A.R., G.S., and L.E.F. wrote the manuscript. E.S., S.M.K., R.H.J.M., A.R., G.S., and L.E.F. designed the research. E.S., S.M.K., R.H.J.M., A.R., G.S., and L.E.F. performed the research. E.S., S.M.K., and L.E.F. analyzed the data.

1. Therasse, P. *et al.* New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl. Cancer Inst.* **92**, 205–216 (2000).
2. Eisenhauer, E.A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
3. Sargent, D.J. *et al.* Validation of novel imaging methodologies for use as cancer clinical trial end-points. *Eur. J. Cancer* **45**, 290–299 (2009).
4. Choi, H. *et al.* CT evaluation of the response of gastrointestinal stromal tumors after imatinib mesylate treatment: a quantitative analysis correlated with FDG PET findings. *AJR Am. J. Roentgenol.* **183**, 1619–1628 (2004).
5. Choi, H. *et al.* Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J. Clin. Oncol.* **25**, 1753–1759 (2007).
6. Benjamin, R.S. *et al.* We should desist using RECIST, at least in GIST. *J. Clin. Oncol.* **25**, 1760–1764 (2007).
7. Ammari, S. *et al.* Radiological evaluation of response to treatment: application to metastatic renal cancers receiving anti-angiogenic treatment. *Diagn. Interv. Imaging* **95**, 527–539 (2014).
8. Dudeck, O., Zeile, M., Reichardt, P. & Pink, D. Comparison of RECIST and Choi criteria for computed tomographic response evaluation in patients with advanced gastrointestinal stromal tumor treated with sunitinib. *Ann. Oncol.* **22**, 1828–1833 (2011).
9. Faivre, S. *et al.* Changes in tumor density in patients with advanced hepatocellular carcinoma treated with sunitinib. *Clin. Cancer Res.* **17**, 4504–4512 (2011).

10. Nishino, M., Giobbie–Hurder, A., Ramaiya, N.H. & Hodi, F.S. Response assessment in metastatic melanoma treated with ipilimumab and bevacizumab: CT tumor size and density as markers for response and outcome. *J. Immunother. Cancer* **2**, 40 (2014).
11. Shady, W. *et al.* Surrogate imaging biomarkers of response of colorectal liver metastases after salvage radioembolization using 90Y-loaded resin microspheres. *AJR Am. J. Roentgenol.* **207**, 661–670 (2016).
12. Taieb, S. *et al.* Comparison of response evaluation criteria in solid tumours and Choi criteria for response evaluation in patients with advanced soft tissue sarcoma treated with trabectedin: a retrospective analysis. *Eur. J. Cancer* **51**, 202–209 (2015).
13. Tirumani, S.H., Shinagare, A.B., O'Neill, A.C., Nishino, M., Rosenthal, M.H. & Ramaiya, N.H. Accuracy and feasibility of estimated tumour volumetry in primary gastric gastrointestinal stromal tumours: validation using semiautomated technique in 127 patients. *Eur. Radiol.* **26**, 286–295 (2016).
14. Eechoute, K. *et al.* A long-term prospective population pharmacokinetic study on imatinib plasma concentrations in GIST patients. *Clin. Cancer Res.* **18**, 5780–5787 (2012).
15. Schiavon, G. *et al.* The effect of baseline morphology and its change during treatment on the accuracy of Response Evaluation Criteria in Solid Tumours in assessment of liver metastases. *Eur. J. Cancer* **50**, 972–980 (2014).
16. Nishino, M. *et al.* CT tumor volume measurement in advanced non-small-cell lung cancer: performance characteristics of an emerging clinical tool. *Acad. Radiol.* **18**, 54–62 (2011).
17. Warren, K.E., Patronas, N., Aikin, A.A., Albert, P.S. & Balis, F.M. Comparison of one-, two-, and three-dimensional measurements of childhood brain tumors. *J. Natl. Cancer Inst.* **93**, 1401–1405 (2001).
18. Aghighi, M. *et al.* Three-dimensional radiologic assessment of chemotherapy response in Ewing sarcoma can be used to predict clinical outcome. *Radiology* **280**, 905–915 (2016).
19. Zhao, B., Schwartz, L.H., Moskowitz, C.S., Ginsberg, M.S., Rizvi, N.A. & Kris, M.G. Lung cancer: computerized quantification of tumor response–initial results. *Radiology* **241**, 892–898 (2006).
20. Nishino, M. *et al.* Tumor volume decrease at 8 weeks is associated with longer survival in EGFR-mutant advanced non-small-cell lung cancer patients treated with EGFR TKI. *J. Thorac. Oncol.* **8**, 1059–1068 (2013).
21. Nishino, M. *et al.* Volumetric tumor response and progression in EGFR-mutant NSCLC patients treated with erlotinib or gefitinib. *Acad. Radiol.* **23**, 329–336 (2016).
22. Bender, B.C., Schindler, E. & Friberg, L.E. Population pharmacokinetic-pharmacodynamic modelling in oncology: a tool for predicting clinical response. *Br. J. Clin. Pharmacol.* **79**, 56–71 (2015).
23. Mould, D.R., Walz, A.C., Lave, T., Gibbs, J.P. & Frame, B. Developing exposure/response models for anticancer drug treatment: special considerations. *CPT Pharmacometrics Syst. Pharmacol.* **4**, e00016 (2015).
24. Venkatakrishnan, K. *et al.* Optimizing oncology therapeutics through quantitative translational and clinical pharmacology: challenges and opportunities. *Clin. Pharmacol. Ther.* **97**, 37–54 (2015).
25. Schindler, E., Amantea, M.A., Karlsson, M.O. & Friberg, L.E. PK-PD modeling of individual lesion FDG-PET response to predict overall survival in patients with sunitinib-treated gastrointestinal stromal tumor. *CPT Pharmacometrics Syst. Pharmacol.* **5**, 173–181 (2016).
26. Schiavon, G. *et al.* Tumor volume as an alternative response measurement for imatinib treated GIST patients. *PLoS One* **7**, e48372 (2012).
27. Ribba, B. *et al.* A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT Pharmacometrics Syst. Pharmacol.* **3**, e113 (2014).
28. Claret, L. *et al.* 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 (2009).
29. Petersson, K.J., Hanze, E., Savic, R.M. & Karlsson, M.O. Semiparametric distributions with estimated shape parameters. *Pharm. Res.* **26**, 2174–2185 (2009).
30. Sharma, A. & Jusko, W.J. Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. *Br. J. Clin. Pharmacol.* **45**, 229–239 (1998).
31. Holford, N. A time to event tutorial for pharmacometricians. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e43 (2013).
32. Zhang, L., Beal, S.L. & Sheiner, L.B. Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *J. Pharmacokinet. Pharmacodyn.* **30**, 387–404 (2003).
33. Beal, S., Sheiner, L.B., Boeckmann, A. & Bauer, R.J. NONMEM user's guides (1989–2009) (Icon Development Solutions, Ellicott City, MD, 2009).
34. Keizer, R.J., Karlsson, M.O. & Hooker, A. Modeling and simulation workbench for NONMEM: tutorial on Pirana, PsN, and Xpose. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e50 (2013).
35. Sheiner, L.B. Analysis of pharmacokinetic data using parametric models. III. Hypothesis tests and confidence intervals. *J. Pharmacokinet. Biopharm.* **14**, 539–555 (1986).
36. Suzuki, K., Kohlbrenner, R., Epstein, M.L., Obajuluwa, A.M., Xu, J. & Hori, M. Computer-aided measurement of liver volumes in CT by means of geodesic active contour segmentation coupled with level-set algorithms. *Med. Phys.* **37**, 2159–2166 (2010).
37. Choi, J.W. *et al.* Small submucosal tumors of the stomach: differentiation of gastric schwannoma from gastrointestinal stromal tumor with CT. *Korean J. Radiol.* **13**, 425–433 (2012).
38. Wardelmann, E. *et al.* Polyclonal evolution of multiple secondary KIT mutations in gastrointestinal stromal tumors under treatment with imatinib mesylate. *Clin. Cancer Res.* **12**, 1743–1749 (2006).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://psp-journal.com)