Novel anticancer therapeutics fail in phase III trials more frequently than do drugs developed in other fields of medicine. Because phase III trials are the most expensive to conduct, this high failure rate disproportionately increases the total costs for development of anticancer therapy. The need to modernize decision making in the phase II–to–phase III transition has been recognized for nearly a decade, and the issue of improving phase II clinical trial design has garnered increasing attention.

Measurement of the effects of treatment for solid tumors has been difficult to standardize. The most widely used method—the Response Evaluation Criteria in Solid Tumors (RECIST)—achieves reproducible results by estimating the quantitative change in tumor burden and then categorizing disease response as complete response, partial response, stable disease, or progressive disease. The cut points to separate disease response into the different categories were intended to minimize miscategorization and were first established when digital imaging was a new technology. At that time, there were few available standard treatments for advanced solid tumors. A categorical system was efficient for screening new agents, especially drugs that were expected to shrink tumors within short time frames. Now that solid tumors have become more treatable, and testing novel agents in combination programs or in subsets of patients has become more common, the inefficiencies of categorical tumor assessment are becoming less sustainable.

Efforts to improve phase II trials are especially timely for renal cell carcinoma (RCC) in which seven new agents in 6 years have been approved by the US Food and Drug Administration for metastatic disease. A promising approach to evaluating new therapies, recently demonstrated in lung and colorectal cancer, is to model the course of disease progression from tumor measurements collected in previous clinical trials. Relying on the sum of the longest dimensions of the measured target lesions rather than RECIST categorization, this approach offers an efficient, adaptable method for evaluating treatments with the current commonly used, primarily computed tomography (CT) imaging data.

Because each disease has different, characteristic, patterns of growth, we have developed a longitudinal tumor growth model for RCC based on measurements collected in the phase III trial that supported regulatory agency approval for sorafenib. We then validated this model with measurements collected in the phase III trial that led to approval of pazopanib. To determine the potential for evaluating therapeutics based on a quantitative assessment of treatment effect, we performed simulations to assess the power of randomized phase II trials to detect different magnitudes of treatment effect above the estimated effect of sorafenib.

Estimation of Renal Cell Carcinoma Treatment Effects From Disease Progression Modeling

ML Maitland, K Wu, MR Sharma, Y Jin, SP Kang, WM Stadler, TG Karrison, MJ Ratain, and RR Bies

To improve future drug development efficiency in renal cell carcinoma (RCC), a disease–progression model was developed with longitudinal tumor size data from a phase III trial of sorafenib in RCC. The best-fit model was externally evaluated on 145 placebo-treated patients in a phase III trial of pazopanib; the model incorporated baseline tumor size, a linear disease-progression component, and an exponential drug effect (DE) parameter. With the model-estimated effect of sorafenib on RCC growth, we calculated the power of randomized phase II trials between sorafenib and hypothetical comparators over a range of effects. A hypothetical comparator with 80% greater DE than sorafenib would have 82% power (one-sided $\alpha = 0.1$) with 50 patients per arm. Model-based quantitation of treatment effect with computed tomography (CT) imaging offers a scaffold on which to develop new, more efficient, phase II trial end points and analytic strategies for RCC.
RESULTS
Longitudinal growth model of RCC
In the Treatment Approaches in Renal Cancer Global Evaluation Trial (TARGET) study sample used to generate the initial model, 36% of patients had one tumor lesion, 20% had two, and 44% had the sum of three lesions available for serial measurement. Initially, linear and exponential models were used to fit the placebo data for estimating tumor growth rate. The linear model described the data better than the two exponential models (Table 1), with lower value of Akaike information criterion and lower objective function value. Models with a drug effect (DE) parameter were explored with the placebo data, but no such parameter could be estimated (lack of convergence), suggesting that there was no average reduction in tumor size in the placebo arm.

The final model simultaneously estimated mean baseline tumor size (BASE), the tumor growth rate parameter (PR), DE in the sorafenib arm, and residual variance using all the available data (both placebo and treatment arms) from the TARGET study. To estimate sorafenib’s effect on RCC growth entailed an exponential treatment-effect parameter. Estimated parameters and their relative SEs (%SEs) are presented in Table 2. Note that NONMEM fits a nonlinear, mixed-effects model incorporating the fixed effects (BASE, PR, DE) as well as random, interindividual variability in these parameters. The %SEs reported from NONMEM are all <25%.

One artifact of hierarchical modeling approaches such as this one is known as Bayesian “shrinkage.” Excessive shrinkage could lead to spurious identification of patient-specific factors associated with tumor progression. Although we did not pursue detailed evaluation of patient-specific factors, we calculated shrinkage in interindividual variability (η), which is reported in Table 2. Because clinical and molecular differences among patients and their tumors, respectively, might lead to distinct observable differences in BASE and progression rate, mixture distributions were explored for BASE and PR parameters in the placebo data set. However, these models were not stable, and model diagnostics reflected significant misspecification with the mixture distribution models. Thus, our model makes the standard assumption that the random effects follow a normal distribution. Finally, bootstrapping was also used to evaluate the uncertainty in parameter estimates, with median values from 1,000 replicates shown (Table 2). The median parameter values resulting from the bootstrap procedure agreed with the estimates from our final model. This suggests that the parameters in the final model were reasonably well determined and the model was stable. From 2,000 bootstrap runs, 1,000 (placebo model) and 947 (combined placebo and sorafenib model) minimized successfully and were included in the bootstrap analysis.

Consideration of informative dropout
In these studies, investigators used their clinical assessments and RECIST criteria to determine when patients had disease progression. This means that the patients’ tumor-measurement data could be censored without additional CT measurements (clinically determined progression) or with solely new, small, difficult-to-measure lesions that indicate RECIST progression but do not contribute significantly to the sum of the longest dimensions of tumor measurements. The probability of dropout (termination of the longitudinal data due to disease progression or censoring) is therefore dependent on this unobserved tumor growth since the most recent RECIST measurements. To determine whether this nonrandom dropout affected the model, we estimated the dropout hazard (βd), which indicated an increased hazard of dropping out with larger tumor size. The last observation was also considered to be related to withdrawal, but the dropout model did not support inclusion of both the last observation and the unobserved tumor size due to their high correlation. With lower objective function value and shrinkage, the unobserved predicted tumor size was included in the dropout model. However, this dropout model had minimal effect on the renal cancer growth model parameter estimates (Table 2). This suggests that the development of new metastatic lesions does not contribute to significant deviation from predictions of this linear growth model. The likelihood of disease progression is captured by the increasing sum of longest dimensions of the original target lesions.

Model performance and external validation
Selected cases of placebo-treated and sorafenib-treated patients demonstrate agreement between predicted and observed data (Figure 1). We performed an external validation of the model with data collected from 145 patients assigned to placebo in the

Table 1 Comparative evaluation of candidate models

| Model                              | AIC | OFV |
|------------------------------------|-----|-----|
| Placebo model                      |     |     |
| T(S)(t) = BASE + PR × t            | 7,872.4 | 7,868.4 |
| T(S)(t) = BASE × e(PR × t)         | 7,996.8 | 7,992.8 |
| T(S)(t) = BASE × t^PR              | 8,256.2 | 8,252.2 |
| Combined placebo and treatment model |     |     |
| T(S)(t) = BASE – DE × t + PR × t   | 18,459.4 | 18,434.3 |
| T(S)(t) = BASE × e(-DE × t) + PR × t | 18,209.5 | 18,203.5 |
| T(S)(t) = BASE × t^DE + PR × t     | 18,534.5 | 18,528.5 |

AIC, Akaike information criterion; BASE, baseline tumor size; DE, drug effect; OFV, objective function value; PR, tumor growth rate parameter; T(S)(t), tumor size at time t.

Table 2 Population parameter estimates for the tumor progression and dropout models

| Parameter      | NONMEM | Bootstrap |
|----------------|--------|-----------|
| Tumor progress model |        |           |
| PR (mm/day)     | 0.158  | 11.6      | 0.158  | 15.2 |
| BASE (mm)       | 62.7   | 2.81      | 62.7   | 2.73 |
| DE (one/day)    | 0.00443| 11.1      | 0.00443| 13.9 |
| Interindividual variability (shrinkage %) |        |           |
| PR (additive)   | 0.22 (17.9) | 21.4 | 0.23 | 27.5 |
| BASE (%)        | 71.5 (3.25) | 5.0 | 71.4 | 4.80 |
| DE (additive)   | 0.005 (46.1) | 23.2 | 0.005 | 27.5 |
| Residual error (shrinkage %) |        |           |
| Proportional (%)| 8.9 (30.8) | 11.7 | 8.9 | 20.23 |
| Dropout model   |        |           |
| Baseline hazard | 0.00806 | 14.6 | — | — |
| Dropout hazard  | 0.00635 | 21.6 | — | — |

BASE, baseline tumor size; DE, drug effect; NONMEM, nonlinear mixed-effects model; PR, tumor growth rate parameter; %SE, relative SE.
Power calculations for DE parameter as an end point in future trials

The DE parameter in the tumor growth model provides a quantitative representation of the effect of drug on tumor growth on a continuous scale. By definition, the placebo takes a value of \( DE = 0 \). In a placebo-controlled trial, the difference in \( DE = DE_{\text{treatment}} - DE_{\text{placebo}} \). For sorafenib in the TARGET trial, \( DE = 0.00443/\text{day} \). One way to interpret the model is that for a patient with a BASE tumor size of 62.7 mm in the placebo arm, on average, the tumor will increase to \( 62.7 + 0.158 \times 30 = 67.4 \text{ mm} \) after 1 month, a 7.5% increase. In the sorafenib arm, after 1 month, the expected size is \( 62.7e^{-0.00443 \times 30} + 0.158 \times 30 = 59.6 \text{ mm} \), a 5% decrease. This may not be as intuitive to the trialist as progression-free survival (PFS) but is a novel quantitative term to reflect the impact of drug in terms of % reduction in tumor size. In the TARGET trial, hazard ratio for disease progression was 0.44 (5.5 months in the sorafenib arm vs. 2.8 months in the placebo arm). A larger DE will result in a lower hazard ratio.

As standard-of-care treatments become available and serve as comparators, clinical trials can be designed to detect evidence of improvement in DE over that achieved with sorafenib. Given the variance in the DE estimate for sorafenib from the 749 subjects for whom data were available, we assessed power for relative improvements in the DE over sorafenib in simulated, randomized, two-arm, phase II trials having 50 patients per arm with a one-sided type I error rate (\( \alpha \)) = 0.1 (Figure 3). The ratio of DE for the comparator as compared with sorafenib ranged from 1.2 to 2.0, for which power ranged from 30 to 91%. Randomized studies (50 patients/arm) had 82% power to detect a significant difference in DE when the comparator had a true DE 1.8 times (or 80% greater than) that of sorafenib and 91% power to detect a significant difference in DE when the comparator had a true DE 2.0 times (or 100% greater than) that of sorafenib.

**DISCUSSION**

We have presented a modeling framework to quantify the magnitude of DE for therapeutics tested in advanced RCC, based on CT imaging measurements collected in phase III trials. The model was first developed with imaging data from the multicenter trial of sorafenib in RCC and then validated with imaging data from the multicenter trial of pazopanib in RCC. Simulations using the

---

**Figure 1** Plots of representative individual patients. (a) Placebo-treated patients and (b) sorafenib-treated patients. Open circles indicate observed tumor sizes for the individual patient, solid lines indicate individual model-predicted tumor sizes, and dashed lines indicate model-predicted tumor sizes for the entire cohort of placebo- (as in a) or sorafenib-treated (as in b) patients. ID, identification number.

**Figure 2** Visual predictive check of the joint tumor size/informative dropout model over 180 days. The open circles represent the observed sum of largest dimensions of target lesions for each placebo-treated patient from the pazopanib trial at each time point. The solid line is the median of the simulated data, and the 90% prediction intervals are encompassed by the dashed lines. Randomized phase III trial of pazopanib. The visual predictive check (Figure 2) depicts the distribution of the sum of the largest dimensions of all target lesions for each patient in the placebo arm over time. The median and 5th and 95th percentiles of the tumor size distributions are captured by the model-based predictions through ~168 days (or four CT imaging evaluations). Importantly, the model captures the population “dip” in tumor size at the 95th percentile over the first, second, and third CT imaging evaluation intervals as the subjects with the largest tumor burdens drop out.

**Figure 3** Placebo-treated patients and (b) patients. ID, identification number.
DE model parameter suggest that a randomized two-arm trial (with 50 patients/arm) would have sufficient power to detect new treatments with at least 80% greater DE than sorafenib. As with any end point, larger clinical trials would be required to detect smaller improvements in DE over sorafenib.

This initial modeling effort to approach RCC progression evaluation on a continuous rather than categorical scale takes into account the full set of longitudinal measurements. This model serves as a modifiable scaffold on which new advances in predictive and prognostic factors could be incorporated to make clinical trials in advanced RCC more efficient. The model that best fit this large data set relies on three fundamental parameters: the baseline burden of disease represented by the sum of the longest dimensions of the measurable lesions on CT imaging (BASE), and the combination of DE and PR, reflecting the growth rate determined by the changes in total tumor burden assessed over time. It is intuitively obvious that differences in “how much tumor you have” and “how fast it is growing” are the primary variables for assessing treatments in advanced disease. Indeed, these are the same parameters used in disease-progression models for colorectal cancer and non–small cell lung cancer. Based on readily collected empirical imaging data, this model does not address the more theoretical, mechanistic considerations of tumor fractions growing and regressing during the course of treatment. However, models of this structure have been robust for application to both first- and second-line treatment settings in other tumor types.

The BASE parameter is typically larger in second-line than in first-line treatment, but the linear progression rate is within the same range. Because these empirical models have been developed only recently, potentially important predictors such as prior therapy response, molecular markers, and comorbid conditions have not yet been incorporated.

For the model in this article, we referred to the exponential decay parameter as DE. In the non–small cell lung cancer model, developed primarily with cytotoxic therapeutics, this term is called the shrinkage rate, and in the colorectal cancer model, developed solely with cytotoxic therapeutics, the term is drug constant cell kill rate. This RCC model has been developed and validated solely with inhibitors of vascular endothelial growth factor receptor 2. These agents may decrease the measured size of tumors in part through killing tumor cells, but slowing the pace of further tumor growth and diminishing intratumoral hydrostatic pressure are other means by which these drugs may decrease size and rate of progression of tumor masses (which can reduce associated symptoms). This therefore justifies the more general term, DE.

As for many different solid tumor types, the advancement of therapy for advanced RCC has entered a new era. Several distinct chemical entities have clear evidence of benefit as compared with placebo, and more recently these have begun to be compared directly against established standard therapies. However, the need to detect evidence of clinically significant benefit for new drugs in disease subsets and to detect benefit of combination treatments with factorially increasing combination possibilities requires new approaches to the conduct of phase II clinical trials. For RCC, the disease-progression model in this study may improve clinical research efficiency in a number of ways. First, as in the conduct of trials using RECIST, the solely required technology is readily available CT imaging. Second, studies to detect meaningful differences in treatment effects among randomly assigned groups are likely to require the fewest patients with a quantitative parameter end point for which intersubject variance can be estimated and incorporated. Third, this and subsequent models can be shared and optimized publicly. For example, if a tumor tissue marker (such as VHL genotype) or novel imaging modality (such as positron emission tomography or CT volume measurements) significantly improves the capacity to assess tumor growth response to treatment, that parameter could be readily incorporated into the model and used by investigators in subsequent trials.

The findings from this modeling study are consistent with the results from other efforts and support innovative approaches to new end points in RCC clinical trials. Simulations of data from completed phase III clinical trials of pazopanib and sunitinib have suggested improved power in phase II trials of RCC for alternative end points to standard response rate or PFS based on RECIST. These studies and our investigation, all conducted on vascular endothelial growth factor–signaling pathway inhibitors, support the use of a change in tumor-size measurement at a fixed early time point, such as 8 weeks, as a more sensitive end point for testing new RCC therapeutics than the conventional RECIST response rate. By obviating the need for the longer-term follow-up of PFS, studies with this end point might also be completed more rapidly and inexpensively. Although a model-based end point of DE promises additional efficiency, it would be simplistic and premature to compare its performance characteristics with PFS. As recently published in this journal, a complete modeling and simulation study, incorporating data from multiple drugs tested in a randomized trial, would be needed to assess the relationship among the model parameter, DE, and PFS.

There is room for improvement on this initial model-building effort. Because of the limitations of the studies in which imaging data were collected, the model describes tumor growth and not patient-survival patterns. However, Heng et al. have suggested...
that PFS and overall survival are well correlated. The variances not accounted for by the model include not only disease heterogeneity but also the flaws in the data collection and transmission process. Although the model structure would be unlikely to change significantly, the parameter estimates would probably be more precise if measurements collected directly from images by a central reviewer were used rather than data extracted from radiology reports and transcribed by research associates. Finally, a recurrent criticism of models derived from sum of longest dimensions measures is that disease progression defined by new lesions is inadequately captured. Our assessment of an informative dropout effect suggests that this might be relevant and worth further development but does not impact the power of this model to detect meaningful treatment effects in a modest-sized cohort of patients randomly assigned to different treatment arms.

Although solid-tumor progression models may share the same general structure across disease types, the estimated parameters of these equations will be disease- and disease subtype–specific. Analyzing data in two of the largest advanced RCC trials ever conducted, this model implies that further advances in development of RCC therapeutics may be made through serial studies that collect CT-imaging measurement data to examine the quantitative estimates of DE. This model enables further refinements and improvements through testing of novel variables and contributions of novel technology. By relying on routinely collected imaging data (CT scans), this model offers a novel opportunity to advance clinical research in RCC.

METHODS

Trials and data. Model development was performed with tumor-measurement data, as collected and reported by investigators from 749 patients (375 in the placebo arm; 374 in the sorafenib arm) enrolled in the phase III TARGET trial.23 According to local standards for adherence to RECIST,13 target lesions were identified and measured by investigators. To represent the typical quality of smaller scale, locally performed, phase II trials, the complete, independently reviewed imaging data were not used. Model validation was performed with centrally reviewed data from the 145 patients assigned to placebo treatment in VEG105192, the multicenter phase III study of pazopanib.24 In both studies, imaging assessments were conducted every 6 weeks for the first 24 weeks and every 8 weeks thereafter. Conduct of these studies was approved by the University of Chicago Institutional Review Board before the project commenced.

Data extraction. To maximize informativeness of investigator-level data, we generated a program script in the R software environment (Supplementary Data online) to extract tumor measurements in a format most consistent with RECIST.15 The script produced a data set for import into the nonlinear-mixed-effects modeling software NONMEM (version VII; level 1, ICON, Ellicott City, MD).34 This data set’s headings included patient identification number, visit, lesion number, lesion size, sum of lesion sizes by RECIST 1.1, sum of lesions, a flag for whether or not a lymph node measurement was incorporated, and the number of the visit. New lesions were identified but not incorporated into the initial model building. Patients were also flagged for clinical progression without a confirmatory CT scan. If a lesion decreased to 0 and was not tracked consistently, that lesion was not included in the patient’s sum of longest diameters, unless it was the only baseline lesion.

Tumor growth and sorafenib effect model development. The longitudinal tumor size data (sum of the longest diameters of the serially tracked target lesions) were fitted with nonlinear mixed-effects modeling. The model was built in three stages. Initially, both linear and exponential models were used to fit the placebo data set for estimating tumor growth rate. The second step was to add the DE data and estimate the DE parameters with the placebo effects fixed. The final model simultaneously estimated BASE, PR, DE, and residual variance on all available data (both placebo and treatment arms) of the TARGET study. The first-order conditional estimation algorithm for calculating the likelihood in NONMEM was used for parameter estimation. The tumor-progression model describes tumor size as a function of time and accounts for the natural growth of the tumor and DE. Linear, exponential, and power functions were used to fit tumor size data as candidate models describing both tumor shrinkage and tumor growth. Mixture distributions were explored to model potential subgroups with different tumor growth and/or responses to treatment. A model using a combination of an exponential-decay (shrinkage) and linear-growth (progression) was developed because it was the best for describing the tumor growth and DE. Model structural selection was guided using the objective function value difference as well as the Akaike information criterion. An α value for significance was set a priori to α < 0.01 (objective function value difference of 6.63 for one degree of freedom). The Akaike information criterion was also used to adjust for degrees of freedom, but the absolute difference was assessed and not statistical significance.

The final model fitted is

\[
\text{TS}(t) = \text{BASE} \cdot e^{-\text{DE} \cdot t} + \text{PR} \cdot t
\] (Eq. 1)

\[\text{TS}(t)\] is the tumor size at time \(t\) for the \(i\)th individual, BASE, is the baseline tumor size, DE, is the exponential tumor shrinkage rate parameter due to the DE, and PR, is the tumor growth rate constant, all for the \(i\)th individual. BASE, PR, and DE, incorporate random deviations for the \(i\)th individual about the respective population mean parameter. For additional details of model development, see Supplementary Methods online.

The final tumor growth model parameter estimates were further evaluated internally using a nonparametric bootstrap. The resampling was performed 1,000 times. The median values and the 2.5th and 97.5th percentiles of the parameter estimates obtained by this analysis were compared with those of the final model.

Model validation and power assessment

External model validation. The model was evaluated with a visual predictive check of the 145 placebo-treated subjects from the VEG105192 multicenter phase III study of pazopanib as described above.24 These placebo data were not included in the initial model development. The visual predictive check was generated using 1,000 simulations from the joint tumor growth and dropout model to assess the predictive performance. A graphical comparison was made between observed data and the model-predicted median and 90% prediction interval.

Power calculations with DE as the end point. Randomized, two-arm (50 patients per arm) phase II trials comparing sorafenib and a hypothetical comparator (with DE as the primary end point) were simulated (with 1,000 replicates) to estimate the power to detect a significant difference between arms (\(\alpha = 0.10\)). Specifically, simulated data for tumor size at 6, 12, 18, and 24 weeks were generated using the BASE and progression rate from the validated placebo model, with DEs ranging from 0% to 100% greater than that for sorafenib (0.00443 (sorafenib effect)), 0.005316, 0.006202, 0.007088, 0.007974, 0.008866 (twice the sorafenib effect)). Simulated data used the same estimates of interindividual variability and residual error as fitted for sorafenib. Population estimates of DE for the two 50-patient arms in each simulated trial (hypothetical comparator vs. sorafenib) were compared using a \(z\)-test, and estimated power was the percentage of trials with a statistically significant difference between the two arms.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/cpt
ACKNOWLEDGMENTS
We are grateful to Sibyl Anderson, Frank Cihon, Chetan Lathia, and Gloria Hofvilaen at Bayer and to Bernard Escudier for sharing the TARGET sorafenib data. James Jiang provided initial assistance with evaluation of investigator-level tumor-measurement data. We also appreciate Ohad Amit and Lini Pandite at GlaxoSmithKline and Cora Sternberg for sharing the pazopanib trial data. Ben Suttle and Peter Bonate were generous in sharing their unpublished findings and manuscript. M.L.M. was supported by National Cancer Institute Mentored Career Development Award grant K23CA124802. M.R.S. and S.P.K. were supported by grant T32GM007019 for training in clinical therapeutics. This research was also supported by the University of Chicago Cancer Research Center grant P30-CA04599 and an ASCO Translational Research Professorship (M.J.R.). R.R.B. is supported through the Indiana Clinical and Translational Sciences Institute (grant UL1RR025761-01) through Eli Lilly and Company, Bayer and GlaxoSmithKline provided tumor-measurement data. Preliminary data from this study were presented, in part, at the American Society of Clinical Pharmacology and Therapeutics Annual Meeting, 2011, in Dallas, TX.

AUTHOR CONTRIBUTIONS
M.L.M., K.W., M.R.S., Y.J., W.M.S., T.G.K., M.J.R., and R.R.B. wrote the manuscript. M.L.M., M.R.S., S.P.K., W.M.S., T.G.K., M.J.R., and R.R.B. designed the research. M.L.M., K.W., M.R.S., Y.J., S.P.K., T.G.K., and R.R.B. performed the research. M.L.M., K.W., M.R.S., Y.J., T.G.K., M.J.R., and R.R.B. analyzed the data. K.W., M.R.S., Y.J., T.G.K., and R.R.B. contributed analytical tools.

CONFLICT OF INTEREST
M.L.M. has performed compensated consulting for Astellas Pharma U.S. and Pfizer. Y.J. is currently employed by and owns stock in Pfizer. S.P.K. is currently employed by Merck & Co., in which he has stock ownership. W.M.S. has performed consulting for Novartis, Pfizer, Roche/Genentech, and Aveo and has been on speakers' bureaus for Bayer and Pfizer. M.J.R. reports a compensated consulting relationship with Mylan. R.R.B. has received support from the Indiana Clinical and Translational Sciences Institute through a gift from Eli Lilly and Company, as well as grants from Merck through the Regenstrief Institute and the National Institute of Child Health and Human Development.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Clinical studies to advance cancer therapeutics depend on objective assessment of treatment effects on the pace of growth of tumors.

WHAT QUESTION DID THIS STUDY ADDRESS?
Longitudinal disease-progression modeling offers a quantitative approach to measure tumor burden over time and offers opportunities to detect evidence of biomarker and treatment effects more quickly with fewer patients than current categorical methods.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE
We have established and validated a new mathematical model of renal cancer progression based on routinely collected data from two phase III clinical trials in this disease.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS
This model serves as a basic scaffold that can be improved over time; we have demonstrated how the model could be implemented in the design and analysis of a prospective randomized phase II trial.

© 2013 American Society for Clinical Pharmacology and Therapeutics
utilizing different anti-cancer agents. Clin. Pharmacol. Ther. 85 (suppl. 1), S58 (2009).

28. Vogelzang, N.J. Treatment options in metastatic renal carcinoma: an embarrassment of riches. J. Clin. Oncol. 24, 1–3 (2006).

29. Maitland, M.L. & Schilsky, R.L. Clinical trials in the era of personalized oncology. CA. Cancer J. Clin. 61, 365–381 (2011).

30. Sharma, M.R. et al. Resampling phase III data to assess phase II trial designs and endpoints. Clin. Cancer Res. 18, 2309–2315 (2012).

31. Stein, W.D. et al. Analyzing the pivotal trial that compared sunitinib and IFN-a in renal cell carcinoma, using a method that assesses tumor regression and growth. Clin. Cancer Res. 18, 2374–2381 (2012).

32. Claret, L., Lu, J.F., Bruno, R., Hsu, C.P., Hei, Y.J. & Sun, Y.N. Simulations using a drug-disease modeling framework and phase II data predict phase III survival outcome in first-line non-small-cell lung cancer. Clin. Pharmacol. Ther. 92, 631–634 (2012).

33. Heng, D.Y. et al. Progression-free survival as a predictor of overall survival in metastatic renal cell carcinoma treated with contemporary targeted therapy. Cancer 117, 2637–2642 (2011).

34. Beal, S.L., Sheiner, L.B. & Boeckmann, A.J. NONMEM Users Guides (Icon Development Solutions, Ellicott City, MD, 2008).

35. Bergstrand, M., Hooker, A.C., Wallin, J.E. & Karlsson, M.O. Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. AAPS J. 13, 143–151 (2011).

36. Post, T.M., Freijer, J.J., Ploeger, B.A. & Danhof, M. Extensions to the visual predictive check to facilitate model performance evaluation. J. Pharmacokinet. Pharmacodyn. 35, 185–202 (2008).

37. Yano, Y., Beal, S.L. & Sheiner, L.B. Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check. J. Pharmacokinet. Pharmacodyn. 28, 171–192 (2001).