See COMMENTARY page 136

Elucidation of Relationship Between Tumor Size and Survival in Non-Small-Cell Lung Cancer Patients Can Aid Early Decision Making in Clinical Drug Development

Y Wang[1], C Sung[1,2], C Dartois[1], R Ramchandani[2], BP Booth[2], E Rock[3] and J Gobburu[1]

Four non-small-cell lung cancer (NSCLC) registration trials were utilized to develop models linking survival to risk factors and changes in tumor size during treatment. The purpose was to leverage existing quantitative knowledge to facilitate future development of anti-NSCLC drugs. Eleven risk factors were screened using a Cox model. A mixed exponential decay and linear growth model was utilized for modeling tumor size. Survival times were described in a parametric model. Eastern Cooperative Oncology Group (ECOG) score and baseline tumor size were consistent prognostic factors of survival. Tumor size was well described by the mixed model. The parametric survival model includes ECOG score, baseline tumor size, and week 8 tumor size change as predictors of survival duration. The change in tumor size at week 8 allows early assessment of the activity of an experimental regimen. The survival model and the tumor model will be beneficial for early screening of candidate drugs, simulating NSCLC trials, and optimizing trial designs.

More than 170,000 new cases of lung cancer are diagnosed every year in the United States.[1] Lung cancer accounted for the highest cancer-related death rate during the past decade, exceeding the combined rates of colon, breast, and prostate cancers. Despite the existence of a significant unmet medical need for new cancer treatments, oncology drugs have one of the lowest rates of successful drug development, at only 5% (ref. 2). Even compounds that reach the phase III development stage have a failure rate of ~60%.

The Critical Path Initiative of the US Food and Drug Administration calls for leveraging existing knowledge from clinical data through the use of quantitative modeling to improve the drug development process. We therefore initiated this project to quantify the relationship between early tumor size reduction and patient survival, using the combined clinical data submitted to the Food and Drug Administration by multiple pharmaceutical companies. From the various cancer types, we selected non-small-cell lung cancer (NSCLC) because it is associated with a high death rate in the clinical trials, this being a required element for developing reliable survival models. Here we show how data mining of four drug registration trials for NSCLC agents enabled us to develop a pharmacostatistical model that links time of death to risk factors and the percentage change in tumor size at 8 weeks after start of therapy. The model can facilitate clinical screening of novel compounds and provides a tool for drug developers to use in performing clinical trial simulations to improve the design of future trials. This model will also be of great value in addressing the limitation inherent in the unknown relationship between mean tumor size change and patient benefit in a phase II trial that used continuous change in tumor size as the primary end point.[3]

RESULTS

Baseline risk factors that influence survival

Among the 11 risk factors evaluated, Eastern Cooperative Oncology Group (ECOG) performance score and baseline tumor size were consistent significant predictors of time to death across the nine treatments (**Table 1**). These two risk factors were retained for developing the survival model. Although lactate dehydrogenase level was found to be a significant factor in all the treatments for which these data were available, it was not included in our model because the erlotinib and

[1]Division of Pharmacometrics, Office of Clinical Pharmacology, US Food and Drug Administration, Silver Spring, Maryland, USA; [2]Division of Clinical Pharmacology 5, Office of Clinical Pharmacology, US Food and Drug Administration, Silver Spring, Maryland, USA; [3]Division of Oncology, Office of New Drugs, US Food and Drug Administration, Silver Spring, Maryland, USA. Correspondence: Y Wang (yaning.wang@fda.hhs.gov)

Received 2 December 2008; accepted 28 March 2009; advance online publication 13 May 2009. doi:10.1038/clpt.2009.64

Table 1 Significance of risk factors

Factors	PCB	PC	DC	DCb	VC	PB	ET	DT	PT
ECOG performance status	•	•	•	•	•	•	•	•	•
Baseline tumor size	•	°	•	•	•	•	•	•	•
Lactate dehydrogenase	•	•	•	•	•				
6-month weight loss (<5% vs. ≥5%)	°	•	°	°	°	•	•		
Sex	•	•	°	°	°	°	°	°	°
Prior radiation	°	°	°	°	°	°	°	•	•
Prior surgery	°	•	°	°	°	°	°	°	°
Stage (III or IV)	°	°	°	°	°	°	°	°	°
Age	°	°	°	°	°	°	°	°	°
Number of prior chemotherapy treatments						°	°	°	°
Best response to prior therapy: CR, PR, SD, or PD						°	°	•	°

•, significant; °, not significant.

CR, complete response; DC, docetaxel and cisplatin; DCb, docetaxel and carboplatin; DT, docetaxel; ECOG, Eastern Cooperative Oncology Group; ET, erlotinib; PB, placebo; PC, paclitaxel and carboplatin; PCB, paclitaxel, carboplatin, and bevacizumab; PD, progressive disease; PR, partial response; PT, pemetrexed; SD, stable disease; VC, vinorelbine and cisplatin.

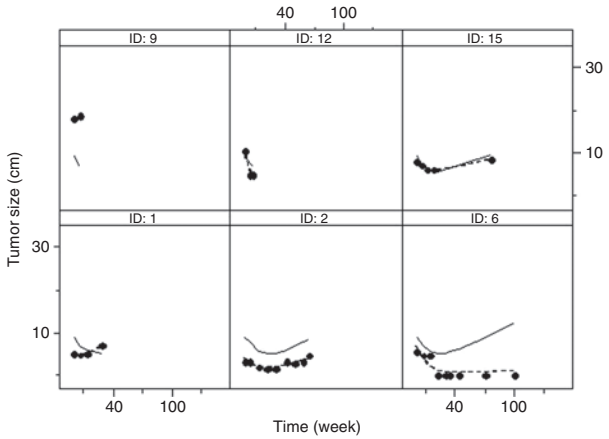

Figure 1 The time course of tumor change for representative individual patients. The symbols represent the observed tumor sizes, the solid line represents the mean tumor size for the overall population, and the broken line represents the individual predicted tumor size.

pemetrexed trials did not have lactate dehydrogenase information. This finding suggests, however, that lactate dehydrogenase data should be collected in future trials and may serve as a stratification factor.

Dynamic model of tumor size

Figure 1 shows individual and population predictions of tumor size dynamics in six representative patients, derived from the mixed exponential decay and linear growth model. The mean profiles of the patient population indicate that the tumors generally undergo shrinkage initially and that this is followed by tumor progression. Although the tumor size in an individual patient can differ substantially from the population prediction, the model is flexible enough to describe various types of individual profiles. Diagnostic plots of the observed and predicted tumor size data for individual patients show that the data are uniformly and closely distributed around the line of identity (**Supplementary Figure S1** online). Table 2 lists the parameter estimates and their estimation precisions. Each treatment was modeled separately so as to obtain the best individual prediction. During model diagnosis of the erlotinib trial, two distinct subpopulations of tumor shrinkage rates were identified. The proportion of patients with a faster shrinkage rate was 12% for placebo and 20% for erlotinib.

Survival model

Baseline tumor size (centered at 8.5 cm), ECOG status (0/1/2/3 as a categorical variable) and percentage tumor reduction from baseline at week 8 (PTR_{wk8}) were found to be the best predictors of time to death (T), as shown in Eq. 1.

$$\log(T) = \alpha_0 + \alpha_1 \times ECOG + \alpha_2 \times (Baseline - 8.5) + \alpha_3 \times PTR_{wk8} + \varepsilon_{TD}, \quad (1)$$

where T is the time to death (day), α_0 is the intercept, α_1, α_2, and α_3 are the slopes for ECOG, centered baseline, and PTR_{wk8}, respectively, and ε_{TD} is the residual variability following a normal distribution with a mean of zero and variance of σ_{TD} (TD; time to death). Because ECOG status is a categorical variable, α_1 represents the difference between ECOG 0 (intercept) and each of the other levels (1/2/3). Log-normal distribution was selected for the survival model (Eq. 1) on the basis of likelihood ratio tests (**Supplementary Table S1** online). An exploratory analysis of the first- and second-line treatments revealed significantly different slopes for PTR_{wk8} and marginally different slopes for ECOG, centered baseline, and intercept. Therefore, different models were fitted to first- and second-line treatments. A separate model was developed for the placebo group of the second-line-treatment population because of the significant differences in several parameter estimates between placebo and other groups, for both included and excluded populations. Parameter values and their estimation precisions are listed in Table 3.

Efforts were made to develop separate survival models for the excluded patients on the basis of baseline risk factors. The median survival times for the excluded patients were consistently shorter than those for the included population. The excluded population contained a higher percentage of patients with higher ECOG scores (Table 4). For clinical trial simulations,

Table 2 Parameter estimates and their precisions (SE) for the tumor size model; interpatient and residual variability expressed as percentages

Treatment	M_BASE (cm)	M_SR (1/week)	M_PR (cm/week)	ω_{BASE}	ω_{SR}	ω_{PR}	σ_{TS}
PCB	9.1 (0.33)	0.06 (0.004)	0.13 (0.02)	59% (0.04)	73% (0.11)	110% (0.39)	15% (0.01)
PC	8 (03)	0.038 (0.01)	0.14 (0.04)	63% (0.06)	98% (0.47)	74% (0.18)	16% (0.03)
DC	8.7 (0.31)	0.052 (0.01)	0.16 (0.02)	57% (0.07)	74% (0.17)	74% (0.14)	20% (0.02)
DCb	9.2 (0.38)	0.047 (0.005)	0.16 (0.02)	64% (0.09)	77% (0.19)	67% (0.12)	18% (0.02)
VC	8.5 (0.28)	0.063 (0.01)	0.17 (0.02)	50% (0.04)	75% (0.18)	92% (0.34)	16% (0.02)
DT	8.5 (0.82)	0.033 (0.01)	0.13 (0.02)	77% (0.23)	190% (1.79)	14% (1.02)	26% (0.06)
PT	7.4 (0.47)	0.023 (0.01)	0.25 (0.05)	70% (0.11)	270% (1.46)	49% (0.18)	14% (0.02)
PB[a]	8.6 (0.44)	0.0047 slow (0.001) 0.13 fast (0.004)	0.20 (0.02)	66% (0.08)	430% (1.63)	60% (0.56)	9% (0.02)
ET[a]	8.4 (0.32)	0.0045 slow (0.001) 0.11 fast (0.05)	0.058 (0.02)	67% (0.09)	280% (0.53)	80% (0.34)	15% (0.02)

DC, docetaxel and cisplatin; DCb, docetaxel and carboplatin; DT, docetaxel; ET, erlotinib; PB, placebo; PC, paclitaxel and carboplatin; PCB, paclitaxel, carboplatin, and bevacizumab; PT, pemetrexed; VC, vinorelbine and cisplatin.

[a]Patient tumors in the study arm that included the PB and ET groups had two distinct populations of shrinkage rates; the fast group constituted 12% of the population in PB and 20% in ET.

Table 3 Parameter estimates and precision (SE) for the survival model for each line of treatment

Treatment	α_0	α_1[a]	α_2	α_3	σ_{TD}
Included population					
First-line	5.75 (0.033)	−0.21 (0.038)	−0.029 (0.0035)	1.07 (0.080)	0.71 (0.015)
Second-line (treatment)	5.91 (0.071)	−0.34 (0.079) −0.61 (0.098) −0.62 (0.19)	−0.029 (0.0048)	0.42 (0.075)	0.72 (0.023)
Second-line (placebo)	5.93 (0.15)	−0.51 (0.16) −0.53 (0.19) −0.94 (0.25)	−0.043 (0.0097)	0.38 (0.16)	0.68 (0.04)
Excluded population					
First-line	5.42 (0.091)	−0.51 (0.12)			1.40 (0.048)
Second-line (treatment)	5.86 (0.21)	−1.25 (0.23) −2.15 (0.25) −2.39 (0.31)			1.14 (0.054)
Second-line (placebo)	5.32 (0.34)	−0.68 (0.37) −1.41 (0.39) −1.75 (0.45)			0.94 (0.074)

[a]For first-line treatment, α_1 represents ECOG1–ECOG0; for second-line treatment, α_1 represents ECOG1–ECOG0, ECOG2–ECOG0, and ECOG3–ECOG0 (top to bottom).

Table 4 ECOG distributions for included and excluded populations

	First line		Second line	
ECOG	Included	Excluded	Included	Excluded
0	799 (52%)	254 (45.8%)	157 (17.3%)	41 (11.3%)
1	739 (48%)	300 (54.2%)	575 (63.4%)	187 (51.7%)
2			148 (16.3%)	98 (27.1%)
3			27 (3%)	36 (9.9%)

it is important to model the excluded patients in proportion to their representation in past trials. In the excluded population, only ECOG score could be identified as a significant predictor of survival time across all nine treatments. In addition, in a significant portion of these patients, baseline tumor measurement data were not available. Therefore, a simplified survival model (Eq. 2) including only ECOG as the predictor was fitted to the excluded population.

$$\log(T) = \alpha_0 + \alpha_1 \times \mathrm{ECOG} + \varepsilon_{TD}. \quad (2)$$

The parameter estimates and their estimation precisions are listed in Table 3. The variance–covariance matrix for the parameter estimates is listed in **Supplementary Table S2** online.

The graphic comparison of survival curves with simulated data and those with observed data showed reasonably good predictive ability in the survival model for each treatment (Figure 2). As compared with the observed median survival times (Table 5), the simulated median survival times tend to shrink toward the mean median survival time of pooled treatments, leading to a conservative estimate of the relative survival benefit. This is reflected in the survival benefit and unstratified hazard ratio (Table 5). For erlotinib and bevacizumab, the model captured

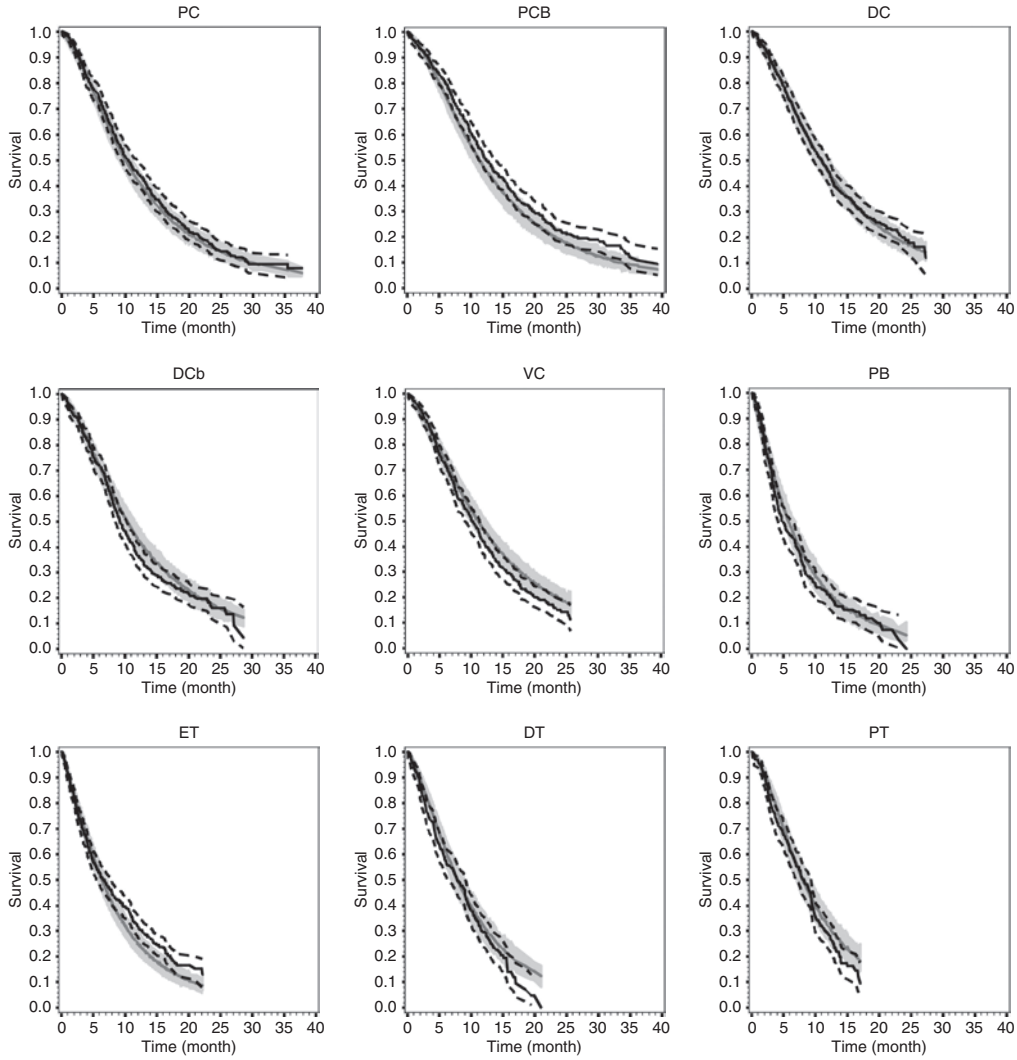

Figure 2 Survival curves for predicted data vs. observed data. The gray line and shaded area represent the mean survival curve and its 95% confidence interval (CI), respectively, relating to predicted data; the solid and broken black lines represent the survival curve and its 95% CI, respectively, relating to observed data.

one-half to two-thirds of the survival benefit. For docetaxel, however, the model did not indicate any survival benefit.

DISCUSSION

Two previous reports from the Food and Drug Administration summarize the regulatory basis and experience using response rate (i.e., change in tumor size) as the basis for accelerated approvals.[4,5] In contrast, the focus of the research described herein is to demonstrate that information in prior registration trials can be leveraged to develop quantitative models that relate patient baseline risk factors and the time course of tumor response to the probability of survival. In addition, the models we have developed can be incorporated into clinical trial simulations to aid early drug development decisions such as molecule screening, mortality trial design, and dose selection.

In agreement with the data published by Brundage *et al.*,[6] we found that ECOG performance status is by itself an important risk factor in NSCLC patients. Previously, Mery *et al.*[7] and Port *et al.*[8] reported hazard ratios for overall survival reflecting the relationship between tumor size category (1–1.9 cm, 2–2.9 cm, 3–3.9 cm, 4–4.9 cm, and 5–6 cm) and time to death in stage 1 or 1A NSCLC patients. According to Mery *et al.*, in a population of 9,191 patients, the hazard ratio gradually increased from 1.27 to 1.95 with increasing baseline tumor size relative to the 1–1.9-cm group. Port *et al.* found that, among 244 stage 1A patients, the hazard ratio was ~1.5 for the group having tumors larger than 2 cm as compared with those with smaller tumors. By contrast, Patz *et al.*[9] found, in a study involving 510 stage 1A NSCLC patients, that there was no correlation between tumor size, whether as a continuous variable or as a categorical variable, and overall survival. Black[10] subsequently offered multiple reasons to explain why the findings of Patz *et al.* might not reflect current understanding of this disease. Although our database had more advanced (stage IIIA/B or IV) NSCLC patients, our results support the conclusions of Mery *et al.* and Port *et al.* This finding in patients with more advanced cancer strengthens the observed relationship between baseline tumor size and prognosis for survival.

Table 5 Comparison of statistics relating to observed and simulated data

	Median survival time in days (95% CI)		Survival benefit in days (95% CI)		Hazard ratio[a] (95% CI)	
Treatment (trial)	Observed	Simulated	Observed	Simulated	Observed	Simulated
PC[b] (1)	315 (285, 357)	328 (289, 371)				
PCB (1)	374 (344, 418)	364 (321, 410)	59	41 (−5, 84)	0.8 (0.69, 0.93)	0.88 (0.78, 1.01)
VC[b] (2)	304 (269, 341)	317 (281, 354)				
DC (2)	331 (292, 368)	309 (274, 347)	27	−7 (−55, 41)	0.89 (0.76, 1.04)	1.04 (0.9, 1.19)
DCb (2)	277 (254, 307)	295 (261, 333)	−27	−22 (−71, 24)	1.03 (0.88, 1.2)	1.07 (0.93, 1.25)
PB[b] (3)	143 (125, 191)	164 (137, 194)				
ET (3)	203 (168, 237)	196 (173, 222)	60	32 (2, 64)	0.76 (0.65, 0.91)	0.83 (0.7, 0.96)
DT[b] (4)	237 (189, 276)	243 (210, 280)				
PT (4)	249 (210, 282)	247 (212, 283)	12	4 (−40, 48)	0.99 (0.82, 1.2)	0.99 (0.85, 1.17)

DC, docetaxel and cisplatin; DCb, docetaxel and carboplatin; DT, docetaxel; ET, erlotinib; PB, placebo; PC, paclitaxel and carboplatin; PCB, paclitaxel, carboplatin, and bevacizumab; PT, pemetrexed; VC, vinorelbine and cisplatin.

[a]Unstratified. [b]Reference arm in each trial.

The time course of tumor growth can be described using several approaches. The Gompertz model is widely applied for modeling tumor growth.[11] Swan[12] describes numerous mathematical models involving differential equations that deal with the dynamic, or time-course, variation of cancer. A more recent model, by Iliadis and Barbolosi,[13] incorporated a cell-loss component into a Gompertz-type growth equation in order to quantify the effect of drug exposure on cell death. Claret *et al.*[14] modified this model in an effort to link drug dose with the sparse tumor size data collected in late-stage clinical trials. A time-varying resistance term was used to link drug dose with tumor growth suppression, whereas another term involving the natural log of tumor size in the Gompertz growth equation was omitted. Tham *et al.*[15] developed an empirical pharmacodynamic model to link drug dose/exposure with tumor growth, using a hypothetical effect compartment. Because the main objective of the tumor model in this project was to interpolate tumor sizes for cases in which observations were not available, we used a more empirical model that includes an exponential component and a linear component to describe the longitudinal tumor size data. This empirical model offered a more stable estimation of tumor size for our database as compared with the other, more complicated models. Patients with a better response to treatment (i.e., decrease in tumor size from baseline) show larger shrinkage rate constants and shallower slopes of linear tumor growth, whereas patients with a poorer treatment response show smaller shrinkage rate constants and steeper slopes of linear tumor growth. Given the sparse data relating to measurements of tumor size in the database and the single-dose regimen in each treatment, this model is flexible enough to describe individual tumor observations well within the time window in the trials (**Figure 1** and **Supplementary Figure S1** online). The limitation of the current model is that it cannot be used to simulate tumor profiles under new dosing regimens because the model does not include dose or exposure. However, if the relevant information, such as longitudinal tumor data from a dose-ranging trial, is available, the tumor size model can be easily expanded to include dose or drug exposure effect on the shrinkage and/or linear growth rate constants, depending on the mechanism of action of the compound.

In order to develop the survival model, tumor size measurements at various time points were explored as potential predictors in addition to those identified as baseline risk factors. Among the time points at which tumor size was evaluated, the week-8 tumor size change was selected because it was consistently identified as a significant predictor across all nine treatments. Importantly, it is also early enough in the treatment to serve as an early biomarker for survival prediction and assist in early drug development decisions. The modeling of survival using tumor size change as a continuous variable offers potentially greater statistical power than treating the tumor size data as a categorical response for comparisons of responses among different treatment groups.[3] The significantly smaller slope for PTR_{wk8} (α_3) for second-line treatment indicates that the survival times for patients under second-line treatment were significantly less sensitive to PTR_{wk8} as compared with those under first-line treatment. The median PTR_{wk8} ranged from 15 to 29% for patients under first-line treatments, whereas it ranged from −14 (tumor size increase) to 12% for patients under second-line treatments. This observation suggests that achieving a median zero tumor growth (stable disease) within 8 weeks is not an adequate end point for a first-line treatment. The placebo group had to be modeled separately, and this may reflect a different tumor shrinkage–survival relationship for certain patients under placebo (best supportive care) as compared with those under active treatments.

The predictive ability of the survival model was evaluated by comparing the survival curves from simulated data with those from the observed data (**Figure 2**). The overlapping of the curves suggests that the model can reasonably reproduce the observed survival curves. Two populations (included and excluded) were simulated and combined to represent the full population for each treatment. The simulated median survival times (**Table 5**) tend to shrink toward the mean median survival time of pooled treatments relative to the observed median survival times. As a result, simulation will generate conservative estimates of the

relative survival benefit. This suggests that the model may be further improved by including more predictors. An improvement in tumor imaging data may increase the predictive power of tumor size for survival.[16] The lack of treatment-related effect in the survival model for the excluded population also contributed to the diminished overall treatment effect. Therefore, we believe that the PTR_{wk8}—measured as the sum of the unidimensional measurements of all measurable lesions based on the Response Evaluation Criteria in Solid Tumors criteria—cannot capture the full survival benefit of a treatment and can serve only as a biomarker for early decision making and as a tool for designing survival trials. However, the conservative estimate of the relative survival benefit between two treatments can prevent overoptimism about the potential survival benefit associated with the use of a new compound, although a slightly better treatment may be discontinued prematurely on the basis of the simulation.

Because survival is an end point that reflects both a drug's effectiveness and its toxicity, an ideal survival model would also take the drug's toxicity into account. In fact, the parameter estimates in Eqs. 1 and 2 may be influenced by each treatment's toxicity, although in our model we included only an effectiveness-related biomarker. The influence of drug toxicity, however, is likely to be minimal, given that the death rate due to this cause is negligible (0.3–3%) as compared with the death rate due to progressive disease (86–94%). It is possible that a drug's toxicity may affect compliance and indirectly influence its effectiveness or a biomarker such as tumor size. Unless an oncology drug's toxicity is severe enough to compete with the disease as a significant factor in causing death, the influence of drug toxicity on overall survival may not be very significant relative to the disease itself.

To the best of our knowledge, our work utilizes the largest database of NSCLC trials to develop a quantitative model of survival benefit using patient risk factors and tumor size data. Similar work for colorectal cancer was performed by Claret *et al.*[14] We show that PTR_{wk8}, along with baseline tumor and ECOG status, can be useful for predicting survival outcome, and this enables drug development decisions to be made earlier. Therefore, a scheduled visit at week 8 with computed tomography imaging for tumor size measurements can provide a critical signal for drug effect. Such information can then be utilized to decide whether to develop that particular drug molecule further. In addition, if a molecule is selected for further development, pivotal trial designs can be optimized using clinical trial simulations. Specifically, we propose the use of baseline tumor size, ECOG status, and PTR_{wk8} in each patient from a phase II trial to simulate the survival benefit based on our model. These simulations should build in 20–30% patients with missing post-baseline measurements, although we urge that every effort be made to collect all relevant data.

The criteria and end points for screening drug molecules are highly variable across investigators. Typically, tumor response rates, time to progression, progression-free survival, and overall survival are reported for phase II studies. Some of these studies are single-arm trials. For example, Monnerat *et al.* reported the median survival time for gemcitabine plus pemetrexed to be 10.1 months (95% confidence interval (CI): 7.9–13.0) for 58 chemonaive patients with stage IIIB/IV NSCLC and ECOG performance scores of 0 or 1 (ref. 17). More recently, Gridelli *et al.* reported the median survival times for patients treated with pemetrexed without and with gemcitabine to be 4.7 months (95% CI: 3.2–6.8) and 5.4 months (95% CI: 3.8–7.6), respectively, for 87 chemonaive patients with stage IIIB/IV NSCLC and ECOG performance scores of 0–2 (ref. 18). Herbst *et al.* found the hazard ratio of progression-free survival for treatment with bevacizumab and chemotherapy was 0.66 (95% CI: 0.38–1.16), and for treatment with bevacizumab and erlotinib, it was 0.72 (95% CI: 0.42–1.23), both in comparison with the standard chemotherapy arm.[19] These estimates of median survival time and hazard ratios are consistently highly uncertain (i.e., wide CIs), mainly because the trials have small numbers of patients and short durations. The lack of reasonable confidence in the predicted outcome hinders decision making for pivotal trial design, and we believe that this contributes to the high failure rate of drugs in phase III trials.[2] The survival model we report here, on the other hand, might reduce the failure rate of drugs in phase III trials by selecting for further development only those candidate compounds that have highly significant survival benefit, on the basis of a conservative estimate of such survival benefit. However, this needs to be evaluated prospectively in future trials. We encourage pharmaceutical sponsors and academic investigators to prospectively employ our model in designing future trials. Routine application of this model will allow us to evaluate the strengths and weaknesses of the model and identify potential improvements. We also believe that factors affecting the tumor size, such as dose or dosing regimen, can be optimized via the tumor model with dose or exposure information incorporated into the model. Finally, we believe that there is value in applying the approach described in this article to the investigation of other cancers and, more important, to linking tumor size effects with animal and/or *in vitro* data so that this model-based approach can be implemented throughout the drug development process.

METHODS

The analysis database included four randomized clinical trials for NSCLC treatments that were submitted for registration: bevacizumab,[20] docetaxel,[21] erlotinib,[22] and pemetrexed.[23] Trial selection was driven by availability of electronic datasets containing longitudinal tumor size measurements and survival data. Eight active treatments and one placebo group were tested in these trials: paclitaxel, carboplatin and bevacizumab, paclitaxel and carboplatin, docetaxel and cisplatin, docetaxel and carboplatin, vinorelbine and cisplatin, placebo, erlotinib, docetaxel, and pemetrexed. The four trials enrolled and followed up a total of 3,398 patients. Patients received either first-line (bevacizumab, docetaxel) or second-line (erlotinib and pemetrexed) treatments for locally advanced or metastatic NSCLC (stages IIIA/B or IV).

First, baseline risk factors for all 3,398 patients were screened to identify factors that influence survival. Next, a dynamic model was developed to describe tumor size changes over time for individual patients for whom both pre- and post-baseline tumor measurements were available. Subsequently, a survival model was developed to link baseline risk factors and change in tumor size to survival time. Therefore, only patients who had both pre- and post-baseline tumor size data and had survived up to at least 8 weeks (2,445 patients, 72%) were included in the survival modeling

(included population). There were 485 patients with no baseline measurements, 404 patients with no post-baseline measurements, 38 patients with unknown ECOG status, and 350 patients who did not survive to 8 weeks. Multiple data variables for a patient could be missing.

Baseline risk factors for survival. The risk factors were selected on the basis of the stratification factors used in the four trials and the potential risk factors for NSCLC identified by Brundage *et al.*[6] These are listed in the first column of **Table 1**. A Cox model was used to screen for consistent significant factors across the nine arms. A stepwise selection method was used with inclusion significance at 0.1 and exclusion significance at 0.05. Risk factors identified by this method were then used in the survival model to account for patient baseline heterogeneity so that the results for the nine arms could be compared.

Tumor size model. Tumor size was recorded as the sum of longest dimensions, using the Response Evaluation Criteria in Solid Tumors criteria.[24] The majority of tumors were measured using computed tomography scan (76–98%) or X-ray (0.8–13%). The remaining tumors were measured by other methods (0.2–6.1%).

The longitudinal tumor size data were analyzed using nonlinear mixed effect models (NONMEM V; Icon Development Solutions, Ellicott City, MD). A model with mixed exponential-decay (shrinkage) and linear-growth (progression) components described the time course of tumor change (Eq. 3).

$$\mathrm{TS}_i(t) = \mathrm{BASE}_i \cdot e^{-\mathrm{SR}_i \cdot t} + \mathrm{PR}_i \cdot t, \tag{3}$$

where $\mathrm{TS}_i(t)$ is the tumor size at time t for the ith individual, BASE_i is the baseline tumor size, SR_i is the exponential tumor shrinkage rate constant, and PR_i is the linear tumor progression rate. The exponential tumor shrinkage component characterizes the treatment effect on tumor shrinkage over time that would result in reducing the tumor size asymptotically toward zero. The rate constant for the shrinkage was restricted to be non-negative. The linear growth component is an approximation of tumor growth under a specific treatment. Therefore, in our model, the slope of linear growth is also treatment dependent. This parameter was also restricted to be non-negative.

Random variability was attributed to two sources: interpatient variability and residual variability. The residual variability captures error caused by model mis-specification and/or in tumor measurements. The unique shrinkage and progression rates (i.e., interpatient variability) of each patient were taken into account, and the population was assumed to follow a log-normal distribution (Eq. 4).

$$\mathrm{BASE}_i = \mathrm{M_BASE} \cdot \exp(\eta_i), \tag{4}$$

where M_BASE is the population median baseline tumor size and η_i is the difference between the individual and population median baseline values on a log scale that is assumed to follow a normal distribution with a mean of zero and variance of ω_{BASE}^2 (interpatient variability, $\mathrm{BPV}_{\mathrm{BASE}}$). The individual SR and PR parameters were also described using similar equations. An exponential error model (Eq. 5) was used for residual variability.

$$\mathrm{TSO}_i(t) = \mathrm{TS}_i(t) \cdot \exp(\varepsilon_i), \tag{5}$$

where $\mathrm{TSO}_i(t)$ is the observed tumor size at time t for the ith individual, $\mathrm{TS}_i(t)$ is the expected tumor size at time t for the ith individual, and ε_i is the difference between the observed and expected values on a log scale. ε_i is assumed to follow a normal distribution with a mean of zero and variance of σ_{TS}^2 (TS: tumor size).

The main purpose of the tumor model was to predict (interpolate) tumor sizes in cases for which observations were not available. Therefore, the main diagnostic criterion in the model was the match between individual predicted values and observed data. When necessary, a mixture model was used to statistically separate the patient population into subgroups so as to arrive at a better individual prediction.

Survival model. Survival modeling comprised two steps: model building and model evaluation. In order to determine the relationship between tumor size change and time to death, the individual tumor size model was used to predict tumor size change at early time points, e.g., 4, 6, and 8 weeks. Predicted tumor size changes were then used to develop a survival model, taking the identified baseline risk factors into consideration. Because the main purpose of our research was to establish a quantitative link between tumor size and survival for use in the design of future trials, a parametric survival model was utilized to relate risk factors and tumor size change with time to death. Several widely used survival functions, including exponential, log normal, Weibull, and gamma, were tested. Likelihood ratio tests and diagnostic plots were applied for selection of an appropriate survival function. This model was fitted to each treatment arm separately in order to assess whether it would be appropriate to pool these arms. Final parameters were updated with a pooled analysis when necessary. A similar approach was applied to derive a survival model based on baseline risk factors for those who were excluded from the tumor size–related survival model (excluded population). SAS version 9.1 (SAS Institute, Cary, NC) was used for these analyses.

Evaluation of the final pooled survival model was conducted through internal evaluation. One thousand sets of parameters were simulated based on the parameter estimates and their variance–covariance matrix from the final pooled models for included and excluded populations, respectively. Then, 1,000 replicates of survival data were simulated for each treatment on the basis of the simulated parameters, the baseline covariates, and $\mathrm{PTR}_{\mathrm{wk8}}$ observed in the trials. The two simulated populations (included and excluded) were combined to form the full population. Median survival curves and 95% CIs were constructed across the 1,000 replicates. The simulated survival curves were compared to the observed survival curves (median and 95% CI). In addition to the graphic comparison of survival curves, different statistical estimates, such as median survival time for each arm, survival benefit, and unstratified hazard ratio relative to the comparator in each trial, were calculated and compared with the observed ones.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/cpt

ACKNOWLEDGMENTS

This research was made possible in part by funding provided by the Office of Women's Health, Food and Drug Administration (FDA). We thank Mark Rothmann for valuable statistical input that greatly improved the manuscript. The views expressed in this article are those of the authors and do not necessarily reflect the official views of the FDA.

CONFLICT OF INTEREST

The authors had no conflict of interest when they were working on this project. However, during the preparation of this manuscript, Dr Sung was working for Rainbow Pharma Consulting in Singapore; Dr Dartois for Novartis in New Jersey; and Dr Rock for GlaxoSmithKline in North Carolina.

© 2009 American Society for Clinical Pharmacology and Therapeutics

1. American Cancer Society. Cancer Facts & Figures 2006 <http://www.cancer.org/downloads/STT/CAFF2006PWSecured.pdf> (2006).
2. Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **3**, 711–715 (2004).
3. Karrison, T.G., Maitland, M.L., Stadler, W.M. & Ratain, M.J. Design of phase II cancer trials using a continuous endpoint of change in tumor size: application to a study of sorafenib and erlotinib in non-small-cell lung cancer. *J. Natl. Cancer Inst.* **99**, 1455–1461 (2007).
4. Johnson, J.R., Williams, G. & Pazdur, R. End points and United States Food and Drug Administration approval of oncology drugs. *J. Clin. Oncol.* **21**, 1404–1411 (2003).
5. Dagher, R., Johnson, J., Williams, G., Keegan, P. & Pazdur, R. Accelerated approval of oncology products: a decade of experience. *J. Natl. Cancer Inst.* **96**, 1500–1509 (2004).
6. Brundage, M.D., Davies, D. & Mackillop, W.J. Prognostic factors in non-small cell lung cancer: a decade of progress. *Chest* **122**, 1037–1057 (2002).
7. Mery, C.M. *et al.* Diameter of non-small cell lung cancer correlates with long-term survival: implications for T stage. *Chest* **128**, 3255–3260 (2005).
8. Port, J.L., Kent, M.S., Korst, R.J., Libby, D., Pasmantier, M. & Altorki, N.K. Tumor size predicts survival within stage IA non-small cell lung cancer. *Chest* **124**, 1828–1833 (2003).

9. Patz, E.F. Jr., Rossi, S., Harpole, D.H. Jr., Herndon, J.E. & Goodman, P.C. Correlation of tumor size and survival in patients with stage IA non-small cell lung cancer. *Chest* **117**, 1568–1571 (2000).
10. Black, W.C. Unexpected observations on tumor size and survival in stage IA non-small cell lung cancer. *Chest* **117**, 1532–1534 (2000).
11. Norton, L., Simon, R., Brereton, H.D. & Bogden, A.E. Predicting the course of Gompertzian growth. *Nature* **264**, 542–545 (1976).
12. Swan, G.W. *Applications of Optimal Control Theory in Biomedicine*. (Marcel Dekker, New York, 1984).
13. Iliadis, A. & Barbolosi, D. Optimizing drug regimens in cancer chemotherapy by an efficacy-toxicity mathematical model. *Comput. Biomed. Res*. **33**, 211–226 (2000).
14. Claret, L. *et al*. Model-based predictions of expected anti-tumor response and survival in Phase III studies based on phase II data of an investigational agent. *Proc. ASCO* **24**, 307s (2006, Abs 6025).
15. Tham, L.S. *et al.* A Pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. *Clin. Cancer Res*. **14**, 4213–4218 (2008).
16. Petrick, N., Brown, D., Suleiman, O. & Myers, K. Imaging as a tumor biomarker in oncology drug trials for lung cancer: the FDA perspective. *Clin. Pharmacol. Ther*. **84**, 523–525 (2008).
17. Monnerat, C. *et al*. Phase II study of pemetrexed-gemcitabine combination in patients with advanced-stage non-small cell lung cancer. *Clin. Cancer Res*. **10**, 5439–5446 (2004).
18. Gridelli, C. *et al*. Single-agent pemetrexed or sequential pemetrexed/gemcitabine as front-line treatment of advanced non-small cell lung cancer in elderly patients or patients ineligible for platinum-based chemotherapy: a multicenter, randomized, phase II trial. *J. Thorac. Oncol*. **2**, 221–229 (2007).
19. Herbst, R.S. *et al*. Phase II study of efficacy and safety of bevacizumab in combination with chemotherapy or erlotinib compared with chemotherapy alone for treatment of recurrent or refractory non small-cell lung cancer. *J. Clin. Oncol*. **25**, 4743–4750 (2007).
20. Sandler, A.B. *et al*. Paclitaxel-carboplatin alone or with bevacizumab for non-small cell lung cancer. *N. Engl. J. Med*. **355**, 2542–2550 (2006).
21. Fossella, F. *et al*. Randomized, multinational, phase III study of docetaxel plus platinum combinations versus vinorelbine plus cisplatin for advanced non-small-cell lung cancer: the TAX 326 study group. *J. Clin. Oncol*. **21**, 3016–3024 (2003).
22. Shepherd, F.A. *et al*. Erlotinib in previously treated non-small-cell lung cancer. *N. Engl. J. Med*. **353**, 123–132 (2005).
23. Hanna, N. *et al*. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. *J. Clin. Oncol*. **22**, 1589–1597 (2004).
24. 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).