JOURNAL OF INTERFERON & CYTOKINE RESEARCH
Volume 31, Number 7, 2011

DOI: 10.1089/jir.2010.0142

Placental Growth Factor and Soluble c-Kit Receptor Dynamics Characterize the Cytokine Signature of Imatinib in Prostate Cancer and Bone Metastases

Paul Mathew,[1,*] Sijin Wen,[2] Satoshi Morita,[3] and Peter F. Thall[2]

To assess the hypothesis that the dynamics of plasma angiogenic and inflammatory cytokines after docetaxel chemotherapy with or without the c-kit/abl/platelet-derived growth factor receptor (PDGFR) inhibitor imatinib mesylate for prostate cancer are associated with outcome, the kinetics of 17 plasma cytokines before versus after chemotherapy were assessed and associations with progression-free survival (PFS) examined. After adjusting for multiple tests, significantly different declines in placental growth factor (PIGF), soluble vascular endothelial growth factor receptor-1 (VEGFR1), VEGF, and soluble c-kit were observed with docetaxel plus imatinib ($n = 41$) compared to docetaxel alone ($n = 47$). Based on a piecewise linear regression model for change in concentration of each cytokine as a function of the probability of change in p-PDGFR *in vivo*, only the dynamics of PIGF ($P < 0.0001$) and soluble c-kit ($P < 0.0001$) differed with imatinib therapy. In a Bayesian log-normal regression model for PFS, a rise in human matrix metalloproteinase 9 after docetaxel alone associated with a longer PFS. Distinct plasma angiogenic cytokines are modified by imatinib and partitioned by *in vivo* p-PDGFR dynamics after docetaxel chemotherapy for metastatic prostate cancer. Plasma PIGF and soluble c-kit kinetics are candidate biomarkers of imatinib effect. The predictive value of human matrix metalloproteinase 9 kinetics for docetaxel efficacy requires prospective validation.

Introduction

Improved outcomes with docetaxel chemotherapy for advanced castration-resistant prostate cancer are being sought with novel combinations that target putative mechanisms of disease progression and drug resistance. Preclinical modeling indicated that the platelet-derived growth factor and its receptor (PDGFR) were upregulated in prostate cancer cells proliferating within the bone microenvironment (Uehara and others 2003). The PDGFR was observed to be upregulated in endothelial cells of vasculature specifically associated with PDGF-expressing tumor, and the PDGFR inhibitor imatinib potentiated taxane efficacy via enhanced endothelial apoptosis, an antivascular effect (Uehara and others 2003; Kim and others 2006).

Contrary to preclinical estimates, a randomized controlled study that compared the efficacy of imatinib in combination with docetaxel versus docetaxel alone in men with castration-resistant prostate cancer and bone metastases showed no added benefit with imatinib (Mathew and others 2007). Unexpectedly, *in vivo* pharmacodynamic monitoring of PDGFR inhibition showed that, within the docetaxel arm, an increased probability of PDGFR activation in peripheral blood leucocytes correlated with improved progression-free survival (PFS) and overall survival (OS) (Mathew and others 2008). Rising plasma PDGF levels were associated with a decreased probability of PDGFR activation and inferior PFS (Mathew and others 2008). While the fundamental biological implications of these observations are yet to be determined, these partitioned outcomes were not equally detected in the docetaxel–imatinib combination arm.

To further explore the dynamic signature of plasma cytokines and their prognostic impact after docetaxel chemotherapy, a panel of 17 additional angiogenic and inflammatory cytokines was constructed. Individual cytokine kinetics between baseline (BL) and after docetaxel exposure, modulation by concurrent PDGF inhibitor therapy, and association with PFS outcomes were studied.

Methods

Patients

One hundred sixteen men were enrolled to a randomized study of docetaxel with placebo or imatinib for metastatic castration-resistant prostate cancer and bone metastases (Mathew and others 2007). Of these, 88 paired plasma samples

Departments of [1]Genitourinary Medical Oncology and [2]Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, Texas.
[3]Department of Biostatistics and Epidemiology, Yokahama City University Medical Center, Yokahama, Japan.
*Present address: Department of Hematology–Oncology, Tufts Medical Center, Boston, Massachusetts.

at BL and 6 weeks later after one cycle of weekly docetaxel-based therapy at cycle 2 day 1 (C2D1) were available.

Multiplex cytokine assay

Plasma levels of all analytes described here were subsequently analyzed in duplicates using a multiplex platform, Meso Scale Discovery (MSD) (Gaithersburg, MD). The analytes were soluble c-kit receptor (c-kit), soluble vascular endothelial growth factor receptor-2 (sVEGFR2, KDR), fibroblast growth factor, VEGF, sVEGFR1, placental growth factor (PIGF), interleukin (IL)2, IL8, IL12p70, IL10, granulocyte macrophage-colony stimulating factor, interferon-γ, IL6, IL10, tumor necrosis factor-α, transforming growth factor-β, and matrix metalloproteinase-(MMP)-9. All reagents were provided with the MSD kits and tests conducted according to the manufacturer's instructions.

Statistical methods

Numerical variables were summarized using means and standard deviations, with association between pairs of variables estimated by Pearson's correlation coefficient (Snedecor and Cochran 1980). The Wilcoxon signed rank test was used for 2 sample comparisons of numerical variables (Hollander and Wolfe 1979), applying the Bonferroni P value correction for multiple tests (Snedecor and Cochran 1980). For each cytokine, the Bayesian regression model and method of Morita and others (2010) were employed to evaluate the effects of change in the cytokine level from BL to C2D1 on PFS time while accounting for the effects of hemoglobin, change in prostate-specific antigen (PSA), and change in p-PDGFR. For each patient, because p-PDGFR was measured in ~2,000 cells both at BL and at C2D1, the within-patient BL and C2D1 distributions of p-PDGFR could be estimated very reliably. Because both the BL and C2D1 distributions of p-PDGFR were clearly bimodal for all patients, the within-patient change in p-PDGFR could not be summarized usefully as the difference between the C2D1 and BL sample means. Rather, a mixture model accounting for the observed bimodality first was fit and used to estimate the differences between the right modes, denoted by δ_{Ri}, and the differences between the left modes, denoted by δ_{Li}, for the within-patient C2D1-versus-BL distributions of p-PDGFR, for each patient, $i = 1, \ldots, 88$.

In the Bayesian regression model for PFS (Morita and others 2010), δ_{Ri} was used as a covariate representing change in p-PDGFR from BL to C2D1. This was done because the values of δ_{Ri} were much larger than δ_{Li}, and moreover δ_{Ri} was strongly associated with longer PFS. Based on preliminary goodness-of-fit analyses, it was assumed that the logarithm of PFS time was normally distributed, equivalently, that PFS was lognormal. The linear component of the lognormal regression model is the mean of log(PFS time), defined as follows. For patient i and cytokine $j = 1, \ldots, 17$, denote the (BL, C2D1) cytokine values by (X_{ij}, Y_{ij}), the difference between the log-transformed cytokine values by $W_{ij} = \log(Y_{ij}) - \log(X_{ij})$, $Z_{1i} = 1$ if treatment was docetaxel+imatinib (DI) and 0 if docetaxel+placebo (DP), Z_{2i} = Hb at BL, and Z_{3i} = change in PSA from BL to C2D1. For cytokine j and patient i, the linear component was assumed to be

$$\begin{aligned}\eta_j = {} & \beta_0 + \beta_1 Z_{1i} + \{\beta_2 Z_{1i} + \beta_3(1 - Z_{1i})\} Z_{2i} \\ & + \{\beta_4 Z_{1i} + \beta_5(1 - Z_{1i})\} Z_{3i} \\ & + \{\beta_6 Z_{1i} + \beta_7(1 - Z_{1i})\} \delta_{Ri} \\ & + \{\beta_8 Z_{1i} + \beta_9(1 - Z_{1i})\} W_{ij}\end{aligned}$$

In terms of their effects on PFS time, the parameters in the linear term may be interpreted as follows:

β_1 = main DI-vs-DP treatment effect
β_2 = effect of baseline Hb in the DI arm
β_3 = effect of baseline Hb in the DP arm
β_4 = effect of change in PSA in the DI arm
β_5 = effect of change in PSA in the DP arm
β_6 = effect of change in p-PDGFR in the DI arm
β_7 = effect of change in p-PDGFR in the DP arm
β_8 = effect of change in cytokine value in the DI arm
β_9 = effect of change in cytokine value in the DP arm

Using the large (n = ~2,000 cells) within-patient p-PDGFR samples taken at BL and at C2D1, the probability of decrease in p-PDGFR after treatment, denoted by Pr(Decr), was estimated very reliably for each patient as a standardized Wilcoxon statistic. Specifically, each patient's Pr(Decr) was computed as the mean over all 0/1 indictors that each BL value of p-PDGFR was larger than each C2D1 value. For each cytokine, the following piecewise linear regression model for the BL to C2D1 change in cytokine value, W_{ij}, as a function of the estimated Pr(Decr) was fit.

$$\begin{aligned}W_{ij} &= b_{0,t} + e_{ij} \text{ if } \Pr(\text{Decr}) \le 0.5 \\ &= b_{0,t} + b_{1,t} * \{\Pr(\text{Decr}) - 0.5\} \\ &\quad + e_{ij} \text{ if } \Pr(\text{Decr}) > 0.5,\end{aligned}$$

for treatment arm t = DI or DP, where e_{ij} denotes normally distributed random measurement error. Under this model, in treatment arm t, on average the BL to C2D1 change in the cytokine value equals the constant $b_{0,t}$ if Pr(Decr) $\le$0.5 and equals the straight line $b_{0,t} + b_{1,t}$ *{Pr(Decr) – 0.5} if Pr(Decr) > 0.5. The cut-off 0.5 was used because Pr(Decr) = 0.5 corresponds to no change in the cytokine from BL to C2D1, whereas Pr(Decr) $\ge$ 0.5 and Pr(Decr) < 0.5 correspond, respectively, to the cytokine going down or up, on average. The piecewise linear form was chosen based on preliminary goodness-of-fit plots of each cytokine change as a function of Pr(Decr). Under the null hypothesis ($b_{0,DP}$, $b_{1,DP}$) = ($b_{0,DI}$, $b_{1,DI}$), the piecewise linear model is the same for the 2 treatment arms. This null hypothesis corresponds to the kinetics of the cytokine, as a function of Pr(Decr), not changing with the addition of imatinib to docetaxel.

Results

The distributions of the 17 plasma angiogenic and inflammatory cytokines at BL and at C2D1 within each treatment arm are summarized in Table 1. These results indicate a significant decline in IL6 and significant increases in PIGF and soluble VEGFR1 in the docetaxel-placebo arm, and a significant decline in soluble c-kit and increase in IL10 in the docetaxel-imatinib arm. Table 2 summarizes changes in cytokine values from BL to C2D1, compared between treatment arms using the Wilcoxon rank sum test. These tests indicate significantly larger declines in PIGF, soluble c-kit,

TABLE 1. MEANS AND STANDARD DEVIATIONS (IN PARENTHESES) OF CYTOKINE VALUES AT BASELINE AND AT COURSE 2 DAY 1 OF CHEMOTHERAPY

	Docetaxel + placebo			*Docetaxel + imatinib*		
Cytokines	*BL*	*C2D1*	*P*	*BL*	*C2D1*	*P*
TGFβ	0.84 (0.22)	0.90 (0.19)	0.009	0.82 (0.22)	0.79 (0.18)	0.586
bFGF	−1.67 (0.24)	−1.67 (0.24)	0.439	−1.65 (0.22)	−1.64 (0.21)	0.881
PIGF	−1.30 (0.09)	−1.20 (0.12)	<0.001[a]	−1.28 (0.09)	−1.35 (0.11)	0.002
sVEGFR1	−0.68 (0.08)	−0.60 (0.10)	<0.001[a]	−0.65 (0.14)	−0.61 (0.26)	0.166
VEGF	−0.77 (0.14)	−0.73 (0.17)	0.05	−0.80 (0.17)	−0.86 (0.16)	0.004
c-kit	0.85 (0.16)	0.86 (0.15)	0.508	0.83 (0.13)	0.70 (0.15)	<0.001[a]
sVEGFR2	1.23 (0.13)	1.24 (0.14)	0.317	1.21 (0.15)	1.19 (0.15)	0.021
hMMP9	1.95 (0.22)	1.99 (0.29)	0.354	1.91 (0.25)	1.83 (0.23)	0.074
GM-CSF	−0.64 (1.14)	−0.68 (1.10)	0.529	−0.47 (0.71)	−0.58 (0.80)	0.05
IFNγ	−0.02 (0.74)	−0.20 (0.77)	0.071	0.13 (0.67)	0.09 (0.89)	0.834
IL10	0.39 (0.92)	0.56 (0.79)	0.019	0.64 (0.67)	0.91 (0.75)	<0.001[a]
IL12p70	0.46 (0.72)	0.50 (0.70)	0.184	0.40 (0.52)	0.39 (0.55)	0.167
IL1β	−0.77 (0.75)	−0.84 (0.73)	0.253	−0.49 (0.64)	−0.58 (0.72)	0.265
IL2	−0.15 (0.55)	−0.27 (0.59)	0.013	0.02 (50)	−0.03 (0.57)	0.677
IL6	0.43 (0.45)	0.06 (0.59)	<0.001[a]	0.57 (0.52)	0.30 (0.54)	0.002
IL8	0.76 (0.20)	0.72 (0.24)	0.068	0.76 (0.18)	0.81 (0.27)	0.178
TNFα	0.90 (0.18)	0.84 (0.19)	0.012	0.97 (0.37)	0.97 (0.32)	0.752

Comparisons of C2D1-versus-BL for each cytokine within each treatment arm were done using the Wilcoxon rank sum test. Using testwise *P* value 0.05 and a Bonferroni adjustment for multiple testing, with 34 tests, a *P* value <0.00147 implies significant change for that cytokine in that treatment arm.

[a]Significant *P* values.

bFGF, basic fibroblast growth factor; BL, baseline; C2D1, cycle 2 day 1; GM-CSF, granulocyte macrophage-colony stimulating factor; hMMP9, human matrix metalloproteinase; IFNγ, interferon gamma; IL, interleukin; PIGF, placental growth factor; sVEGFR, soluble vascular endothelial growth factor receptor-2; TGFβ, transforming growth factor beta; TNFα, tumor necrosis factor alpha.

VEGF, and sVEGFR1 in the docetaxel-imatinib arm compared to the docetaxel-placebo arm, on average. The largest individual quantitative difference in cytokines between the arms was the decline in soluble c-kit in the docetaxel-imatinib arm.

TABLE 2. MEANS AND STANDARD DEVIATIONS (IN PARENTHESES) OF CHANGE FROM BASELINE TO COURSE 2 DAY 1 OF CHEMOTHERAPY FOR EACH CYTOKINE VARIABLE, WITHIN EACH TREATMENT ARM, COMPARED BETWEEN ARMS USING THE WILCOXON RANK SUM TEST

Cytokines	*Docetaxel + placebo*	*Docetaxel + imatinib*	*P*
TGFβ	0.07 (0.23)	−0.03 (0.22)	0.020
bFGF	0.01 (0.28)	0.03 (0.28)	0.847
PIGF	0.12 (0.14)	−0.08 (0.14)	<0.0001[a]
sVEGFR1	0.07 (0.12)	0.04 (0.24)	0.001[a]
VEGF	0.04 (0.13)	−0.06 (0.14)	<0.0001[a]
c-kit	<0.01 (0.08)	−0.14 (0.12)	<0.0001[a]
sVEGFR2	0.01 (0.07)	−0.03 (0.08)	0.017
hMMP9	0.04 (0.25)	−0.08 (0.26)	0.049
GM-CSF	−0.04 (0.99)	−0.08 (0.99)	0.509
IFNγ	−0.20 (0.94)	0.11 (1.24)	0.099
IL10	0.19 (0.51)	0.32 (0.52)	0.137
IL12p70	0.04 (0.22)	−0.01 (0.70)	0.075
IL1β	−0.09 (0.94)	−0.06 (1.07)	0.913
IL2	−0.14 (0.66)	0.01 (0.50)	0.095
IL6	−0.39 (0.49)	−0.27 (0.48)	0.278
IL8	−0.05 (0.20)	0.09 (0.46)	0.053
TNFα	−0.06 (0.18)	0.02 (0.23)	0.042

Using testwise *P* value 0.05 and a Bonferroni adjustment for multiple testing, with 17 tests, a *P* value <0.00294 implies significant change for that cytokine in that treatment arm.

[a]Significant *P* values.

The fitted piecewise linear regression models are summarized in Table 3. For each cytokine, the test of $(b_{0,DP}, b_{1,DP})$ $(b_{0,DI}, b_{1,DI})$ between the 2 treatment groups was performed using an *F* statistic with (2, 84) degrees of freedom. The results indicate that, among the 17 cytokines, the kinetics of 2 specific angiogenic cytokines, PIGF and soluble c-kit, differed significantly between the 2 treatment arms in terms of relationship to *in vivo* p-PDGFR dynamics, as summarized by Pr(Decr). These 2 cytokines were previously identified as among the 4 cytokines decreasing in the docetaxel–imatinib arm compared to the docetaxel–placebo arm (Table 2).

A total of 17 Bayesian log-normal regression models for PFS were fit, one for each cytokine. Because it would be far too cumbersome to tabulate all 17 fitted models, we present only the estimated effects of the C2D1-versus-BL cytokine changes, within each treatment arm, on PFS time. These are the parameters denoted above by β_8 and β_9 in the model linear component. Because parameters are random quantities under a Bayesian model, each parameter has a posterior distribution under the fitted model. For each combination of cytokine and treatment arm, Fig. 1 summarizes the posterior distribution of the parameter in terms of a 95% credible interval. This interval is represented by a vertical line running from the 2.5th percentile up to the 97.5th percentile of the effect's posterior distribution, with the posterior mean represented by an open circle for the DI arm and by a filled circle for the DP arm. Thus, each vertical line summarizes the middle 95% of the effect's posterior distribution. Under the Bayesian model, a line having lower limit near or above the horizontal line at 0 corresponds to a significant increase in PFS as a function of the C2D1-versus-BL cytokine change. For example, a line for β_8 having lower limit 0 would correspond to posterior probability $\Pr(\beta_8 > 0 \mid \text{data}) = 0.975$.

TABLE 3. SUMMARIES OF 17 FITTED REGRESSION MODELS, ONE FOR EACH CYTOKINE

		Docetaxel + placebo		*Docetaxel + imatinib*		*Test for homogeneity between treatment groups*
Cytokine	*Parameter*	*Estimate*	*SE*	*Estimate*	*SE*	*P value*
TGFβ						0.013
	Intercept	0.024	0.034	−0.061	0.041	
	Slope	1.825	0.661	0.555	0.379	
bFGF						0.431
	Intercept	−0.038	0.044	−0.008	0.053	
	Slope	1.826	0.858	0.504	0.492	
PlGF						<0.001[a]
	Intercept	0.131	0.023	−0.084	0.027	
	Slope	−0.336	0.442	0.037	0.253	
sVEGFR1						0.772
	Intercept	0.070	0.030	0.052	0.036	
	Slope	0.036	0.585	−0.228	0.335	
VEGF						0.004
	Intercept	0.032	0.022	−0.067	0.026	
	Slope	0.321	0.421	0.114	0.241	
c-kit						<0.001[a]
	Intercept	0.005	0.017	−0.139	0.020	
	Slope	−0.046	0.321	−0.005	0.184	
sVEGFR2						0.157
	Intercept	0.01	0.012	−0.018	0.015	
	Slope	−0.01	0.235	−0.16	0.135	
hMMP9						0.111
	Intercept	0.041	0.041	−0.095	0.05	
	Slope	−0.097	0.797	0.255	0.457	
GM-CSF						0.122
	Intercept	−0.160	0.159	0.073	0.190	
	Slope	5.201	3.057	−2.396	1.752	
IFNγ						0.630
	Intercept	−0.265	0.178	0.010	0.212	
	Slope	2.894	3.422	1.531	1.962	
IL10						0.246
	Intercept	0.245	0.083	0.358	0.100	
	Slope	−2.28	1.606	−0.552	0.921	
IL12p70						0.474
	Intercept	0.031	0.081	0.111	0.097	
	Slope	0.505	1.558	−1.782	0.893	
IL1β						0.922
	Intercept	−0.112	0.164	−0.034	0.195	
	Slope	0.986	3.148	−0.463	1.805	
IL2						0.475
	Intercept	−0.117	0.097	0.020	0.116	
	Slope	−0.865	1.865	−0.097	1.069	
IL6						0.169
	Intercept	−0.447	0.078	−0.236	0.093	
	Slope	2.393	1.499	−0.464	0.860	
IL8						0.156
	Intercept	−0.062	0.056	0.114	0.067	
	Slope	0.638	1.077	−0.36	0.618	
TNFα						0.129
	Intercept	−0.071	0.033	0.037	0.040	
	Slope	0.348	0.632	−0.227	0.362	

In each model, the change in cytokine value from BL to C2D1 is a piecewise linear function of the estimated Pr(Decr) for p-PDGFR, with different parameters for the 2 treatment groups, where Pr(Decr) is the estimated probability that p-PDGFR decreased from BL to C2D1. For each fitted model, the test for identical intercept and slope parameters in the treatment groups, "homogeneity," is based on an *F*-statistic with (2, 84) degrees of freedom. Using testwise *P* value 0.05 and a Bonferroni adjustment for multiple testing, with 17 tests, a *P* value <0.00294 implies significant heterogeneity between treatment groups, implying different p-PDGFR dynamics with versus without imatinib for that cytokine.

[a]Significant *P* values.

PDGFR, platelet-derived growth factor and its receptor; SE, standard error.

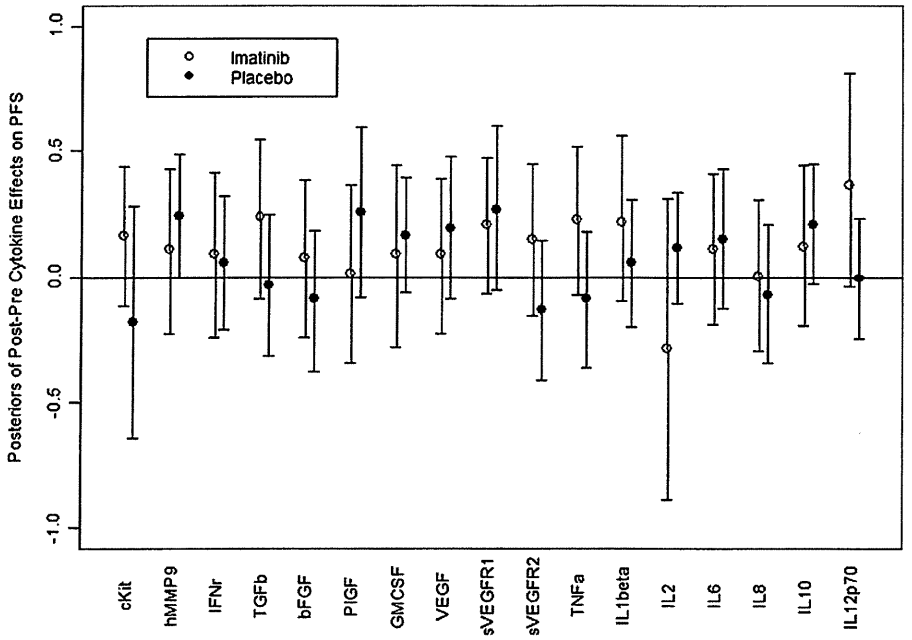

FIG. 1. Estimated posterior effect of each cycle 2 day 1-to-baseline cytokine change on progression-free survival (PFS) the baseline to cycle 2 day 1 change on PFS time for each cytokine within each treatment arm. Each effect was estimated under a Bayesian log-normal regression model, also accounting for the effects of Hb, change in prostate-specific antigen, and change in p-platelet-derived growth factor and its receptor. The posterior distribution of the parameter quantifying the effect of the in terms of a 95% credible interval. This interval is represented by a vertical line running from the 2.5th percentile to the 97.5th percentile of the effect's posterior distribution, with the posterior mean represented by an *open circle* for the docetaxel+ imatinib arm and a *filled circle* for the docetaxel+placebo arm.

This would say that, given the observed data, the posterior probability that the effect of the cytokine's change on PFS is positive equals 0.975, a nominally significant effect. A vertical line with mean at 0 would correspond approximately to posterior probability $\Pr(\beta_8 > 0 \mid \text{data}) = 0.50$, interpreted as the cytokine change having no effect on PFS. Figure 1 shows that, in the DP arm, human MMP9 (hMMP9) had a significant effect, whereas nearly significant effects on PFS were seen for soluble VEGFR1 and IL-10. In the DI arm, a nearly significant effect on PFS was seen for IL-12p70.

Discussion

In this study, the kinetics of 17 angiogenic and inflammatory cytokines in men with metastatic castration-resistant prostate cancer receiving docetaxel with or without the c-kit/abl/PDGFR inhibitor imatinib mesylate were examined. Post-treatment cytokines are significantly modified compared to BL in both treatment arms (Table 1), and several differences vary significantly between both treatment arms (Table 2). Our prior observations had indicated that the status of p-PDGFR activation in peripheral blood leucocytes after docetaxel chemotherapy for castration-resistant prostate cancer associated with PFS and OS (Mathew and others 2008). We then studied the differences in cytokine kinetics between the 2 treatment arms when specifically partitioned by post-treatment *in vivo* p-PDGFR dynamics in peripheral blood leucocytes (Table 3). We find that among these 17 cytokines, PIGF and soluble c-kit dynamics specifically comprise the cytokine signature of imatinib effect after docetaxel chemotherapy.

Decline in soluble c-kit after imatinib therapy has been previously reported in gastrointestinal stromal tumors and has been proposed as a predictive factor for favorable outcome in that disease state (Bono and others 2004, DePrimo and others 2009). In this study, soluble c-kit decline in the imatinib-containing arm was the largest quantitative cytokine difference between the 2 arms. Along with PIGF kinetics, soluble c-kit post-treatment differences retained strong statistical significance when partitioned by p-PDGFR dynamics in peripheral blood leucocytes. These observations may be concordant with the mechanism of action of imatinib as a PDGFR and c-kit inhibitor.

Surprisingly however, in the imatinib arm, increases in soluble c-kit rather than decreases trended toward a favorable PFS profile (Fig. 1) and similarly larger post-treatment values of PIGF and VEGF after docetaxel-alone therapy trended toward an improved PFS. Together, these trends suggest that the cytokine profiles associated with imatinib (c-kit, PIGF, and, to a lesser extent, VEGF declines) compare unfavorably when compared to those generated by docetaxel alone. These findings are also compatible with our previous observations that decreased activation of p-PDGFR in peripheral blood leucocytes after imatinib exposure associated with shorter PFS times (Mathew and others 2008). With the exception of hMMP9 kinetics after docetaxel therapy alone, multivariate analysis of individual cytokine profiles did not yield an independent predictor of outcome. It is conceivable that, with larger numbers of patients, a composite picture of a favorable cytokine signature potentially linked to an *in vivo* mechanism of action of docetaxel may emerge through such cytokine profiling studies.

Declines in the angiogenic cytokines, PIGF, and VEGF after imatinib therapy have not been reported previously. The altered dynamics of these cytokines together with those previously reported with PDGF (Mathew and others 2008) comprise a candidate cytokine signature of imatinib effect in prostate cancer and bone metastases after docetaxel chemotherapy. Formal mechanistic studies will be required to identify the putative link between the regulation of plasma PIGF and VEGF levels and imatinib therapy. It is conceivable that kinetics of these markers may have predictive value in other disease states, hematological and solid neoplasia, in which imatinib has been established as standard therapy, as these circulating cytokines may not be tumor specific.

Before this report, there have been few studies that demonstrate the profile of changes and/or the predictive value of inflammatory and angiogenic cytokine dynamics after docetaxel therapy in prostate cancer. The wide range of

nonhematological toxicities observed with docetaxel, such as peripheral edema or pleural effusions that reflect vascular effects, or fatigue and pneumonitis that reflect proinflammatory effects, are likely to be reflected in plasma cytokine dynamics after treatment. In 2 prior studies, declines in plasma IL6 associated with PSA-declines after docetaxel were reported; however, associations with PFS or OS were not assessed (Domingo-Domenech and others 2006; Ignatoski and others 2009). Our observations do not support a significant association of IL6 decline after docetaxel with PFS (Fig. 1). While significant increases in PlGF and sVEGFR1 and significant decreases in IL6 were observed after docetaxel therapy (Table 1), only an increase in hMMP9 associated with improved PFS (Fig. 1). Elevated hMMP9 expression in prostate cancer has been associated with improved disease-free and OS after prostatectomy for localized prostate cancer (Boxler and others 2010), but a link of plasma MMP9 dynamics with docetaxel efficacy has not been described to our knowledge. These findings suggest the potential predictive value of a cytokine dynamic signature after chemotherapy for prostate cancer, for which larger prospective studies will be required for validation.

Acknowledgments

The authors acknowledge the assistance of Erin Horne and Sherryl Smith (Department of Genitourinary Medical Oncology) for research and administrative support and the M.D. Anderson Cancer Center Immune Monitoring Core Laboratory (IMCL) for assistance with the immunological assays. The IMCL is funded by the M.D. Anderson Cancer Center Support Grant (NCI # CA16672). P.F.T.'s research was partially supported by NCI grant RO1-CA-83932. The authors thank two referees for their detailed and constructive comments.

Author Disclosure Statement

P.M. is on the Speaker Bureau of Sanofi-Aventis. No other disclosures.

References

Bono P, Krause A, von Mehren M, and others. 2004. Serum KIT and KIT ligand levels in patients with gastrointestinal stromal tumors treated with imatinib. Blood 103:2939–2935.

Boxler S, Djonov V, Kessler TM and others. 2010. Matrix metalloproteinases and angiogenic factors: predictors of survival after radical prostatectomy for clinically organ-confined prostate cancer? Am J Pathol 177:2216–2224.

DePrimo SE, Huang X, Blackstein M, Garrett CR, Harmon CS, and others. 2009. Circulating levels of Soluble KIT serve as a biomarker for clinical outcome in gastrointestinal stromal tumor patients receiving sunitinib following imatinib failure. Clin Cancer Res 15:5869–5877.

Domingo-Domenech J, Oliva C, Rovira A and others. 2006. Interleukin-6, a nuclear factor-kB target, predicts resistance to docetaxel in hormone-independent prostate cancer and nuclear factor-kB inhibition by PS-1145 enhances docetaxel antitumor activity. Clin Cancer Res 12:5578–5586.

Hollander M, Wolfe DA. 1979. Introduction to the theory of nonparametric statistics. New York: John Wiley.

Ignatoski KMW, Friedman J, Escara-Wilke J, and others. 2009. Change in markers of bone metabolism with chemotherapy for advanced prostate cancer: interleukin-6 response is a potential early indicator of response to therapy. J Interferon Cytokine Res 29:105–111.

Kim SJ, Uehara H, Yazici S, and others. 2006. Targeting platelet-derived growth factor receptor on endothelial cells of multidrug-resistant prostate cancer. J Natl Cancer Inst 98:783–793.

Mathew P, Thall P, Bucana CD, and others. 2007. Platelet-derived growth factor receptor inhibition and chemotherapy for castration-resistant prostate cancer with bone metastases. Clin Cancer Res 13:5816–5824.

Mathew P, Thall PF, Wen S, and others. 2008. Dynamic change in phosphorylated platelet-derived growth factor receptor in peripheral blood leucocytes following docetaxel therapy predicts progression-free and overall survival in prostate cancer. Br J Cancer 99:1426–1432.

Morita S, Thall PF, Bekele BN, Mathew P. 2010. A Bayesian hierarchical mixture model for platelet derived growth factor receptor phosphorylation to improve estimation of progression-free survival in prostate cancer. J R Stat Soc C 59:19–34.

Snedecor GW, Cochran WG. 1980. Statistical methods, 7th ed. Ames: Iowa State University Press.

Uehara H, Kim SJ, Karashima T, and others. 2003. Effects of blocking platelet-derived growth factor-receptor signaling in a mouse model of experimental prostate cancer bone metastases. J Natl Cancer Inst 95:458–570.

Address correspondence to:
Dr. Paul Mathew
Department of Hematology–Oncology
Tufts Medical Center
800 Washington St., # 245
Boston, MA 02111

E-mail: pmathew@tuftsmedicalcenter.org

Received 17 November 2010/Accepted 17 January 2011

Research Article

Statistics in Medicine

Received 15 January 2010, Accepted 27 May 2010 Published online 18 April 2011 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.3999

Application of the continual reassessment method to a phase I dose-finding trial in Japanese patients: East meets West

Satoshi Morita*†

After cancer-related phase I dose-finding trials are completed in Western countries, further phase I trials are often conducted to determine recommended doses (RDS) for Japanese patients. This may be due to concerns about possible differences in treatment tolerability between Caucasians and Japanese. In most of these, a conventional '3+3' cohort study design is used in making dose escalation decisions, possibly due to its relatively easy implementation. Since its proposal by O'Quigley *et al.* (1990; *Biometrics*, 46:33–48), the continual reassessment method (CRM) has been used increasingly in cancer-related phase I dose-finding studies as an alternative to '3+3' designs. One of the principal advantages of applying a Bayesian CRM may be the utilization of all available prior information to estimate RDS through prior distributions that are assumed for model parameters representing the dose–toxicity relationship. In this paper, we present an application of the Bayesian CRM to a phase I dose-finding study in Japanese patients with advanced breast cancer using an informative prior elicited from clinical investigators. In some settings, it may be appropriate to use an informative prior that reflects the accurate and comprehensive previous knowledge of clinical investigators. On the other hand, for a model-based Bayesian outcome-adaptive clinical trial, it is necessary to establish sufficiently vague priors so that accumulating data dominate decisions as the amount of observed data increases. Thus, we retrospectively investigated the relative strength of the prior using a recently proposed method to compute a prior effective sample size. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: continual reassessment method; dose-finding; phase I trial; prior distribution; prior effective sample size

1. Introduction

After cancer-related phase I dose-finding trials are completed in Western countries, Japanese investigators often conduct trials using the same regimens in Japan to find the optimal doses for Japanese patients. This may be because of concerns about possible differences in treatment tolerability between Caucasians and Japanese. In many cases, recommended doses (RDs) of treatments have been set at higher levels in Caucasians than in Japanese. For example, a phase I study of Taxotere (docetaxel) monotherapy was undertaken in Caucasians to test dose levels from 5 to 115 mg/m^2 [1]. This study identified 100 mg/m^2 as the RD. A subsequent phase I study in Japan tested dose levels from 20 to 90 mg/m^2, and determined that 60 mg/m^2 was the RD for Japanese patients [2].

Japanese clinical investigators develop phase I trial study designs using observed toxicity data and RD levels identified in Western trials as pre-study information. For example, they test a smaller number of dose levels than the original study at doses that account for the RDs in Caucasian patients. In most of these Japanese phase I trials, a conventional '3+3' cohort design is used for making dose escalation decisions, possibly due to its relatively easy implementation and statistical simplicity and the fact that clinical investigators are in general quite familiar with it.

Since its proposal by O'Quigley *et al.* [3], the continual reassessment method (CRM) has been increasingly used in phase I dose-finding studies in cancer patients as an alternative to the '3+3'

Department of Biostatistics and Epidemiology, Yokohama City University Medical Center, 4-57 Urafune-cho, Minami-ku, Yokohama 232-0024, Japan

**Correspondence to: Satoshi Morita, Department of Biostatistics and Epidemiology, Yokohama City University Medical Center, 4-57 Urafune-cho, Minami-ku, Yokohama 232-0024 Japan.*

†E-mail: smorita@urahp.yokohama-cu.ac.jp

design. The CRM, based on a Bayesian parametric model that includes a logistic and a power model [3, 4] is characterized by one or more model parameters representing the dose–toxicity relationship. Although two-parameter models are flexible, they generally require a larger number of patients to estimate two model parameters, e.g. intercept and slope. One-parameter models that analyze one aspect of the dose–toxicity curve (in many cases, the slope) may not be flexible enough to accurately estimate the entire dose–toxicity curve. However, because a one-parameter model in the CRM has proven to be robust in determining a RD [3], it may be reasonable to use a one-parameter model for dose-finding in a cancer phase I trial.

The prior distributions assumed for model parameters are derived from pre-study information and are updated based on accumulated toxicity data observed in consecutive patient cohorts. The prior distribution of the model parameter should reasonably represent clinical investigators' uncertainty about the dose–toxicity relationship before starting the study, sometimes based on historical data from previous clinical studies. A Bayesian approach that formally uses historical/external data to establish such a prior distribution has not yet been fully developed. However, the integration of any available prior information into the estimation of RD levels for Japanese patients may be one of the major advantages of applying Bayesian CRM.

In some settings, it may be appropriate to use an informative prior that reflects the accurate and comprehensive knowledge that clinical investigators already possess. On the other hand, in other cases one may need to avoid excessively informative priors that may unduly influence posterior inferences. In particular, for clinical trials with a model-based Bayesian outcome-adaptive design, it is necessary to establish sufficiently vague priors so that accumulating data dominate decisions as the amount of observed data increases. After completing a Japanese phase I trial, we were concerned about the strength of the established prior distribution relative to the observed data in the trial in which 16 patients were enrolled in total. Thus, we retrospectively investigated the relative strength of the prior using a recently proposed method to compute a prior effective sample size (ESS) [5]. In this paper, we present an application of the CRM to a phase I dose-finding study in Japanese patients with advanced breast cancer using an informative prior elicited from Japanese clinical investigators.

Section 2 provides a motivating example. In Section 3, we describe the application of the CRM to a Japanese phase I study. We discuss establishment of a prior assumed for a dose–toxicity relationship in Section 4. We close with a discussion in Section 5.

2. A motivating example

Although chemotherapy regimens utilizing infusional 5-FU, e.g. the CEF-infu regimen (cyclophosphamide, epirubicin, and infusional 5-FU) [6], have been shown to have high antitumor activity, such regimens require central venous access and pumps. To avoid these inconveniences, a research team from the European Organization for Research and Treatment of Cancer (EORTC) conducted a phase I dose-finding study to develop a new combination regimen substituting the infusional 5-FU in CEF-infu with capecitabine [7]. Capecitabine (Xeloda®) is a novel oral 5-FU prodrug with high single-agent antitumor activity in metastatic breast cancer [8, 9], and also represents an attractive combination partner for the other CEF-infu chemotherapeutic agents [10–12]. The primary objective of the EORTC study was to determine the RD of capecitabine in combination with epirubicin and cyclophosphamide (CEX) in patients with advanced breast cancer. In the EORTC CEX study, four dose levels were planned for capecitabine in combination with fixed doses of epirubicin and CEX (100 and 600 mg/m^2, day 1, every 3 weeks), as summarized in Table I. Capecitabine was escalated from 750 to 1250 mg/m^2 twice daily for three weeks in four dose levels. A conventional '3+3' cohort design was used when making dose escalation decisions. That is, escalation to the next dose level was permitted if zero out of three (0/3) or one out of six (1/6) patients experienced dose-limiting toxicity (DLT). DLT is usually defined as the occurrence of grade 4 hematologic toxicity and grade 3 or 4 non-hematologic toxicity. If more than one patient developed a DLT, the maximum toxic dose (MTD) was reached, and the previous dose level was defined as the RD for phase II studies. In this study, 11 patients received CEX at four dose levels. While defining the MTD, three, three, three, and two patients were entered at dose levels 1, 2, 3, and 4, respectively, as shown in Table I. No DLTs occurred at dose levels 1, 2, and 3. At dose level 4, two out of two patients experienced DLTs. In addition, a high rate of capecitabine treatment modification (interruption and/or reduction) was required at dose level 3. Thus, the EORTC CEX study concluded

Table I. Dose levels of epirubicin and capecitabine studied in the Japanese and EORTC CEX studies and incidence of dose-limiting toxicities (DLTs) observed in the EORTC CEX study. The dose level of cyclophosphamide was fixed at 600 mg/m^2 on day 1 in both studies.

	Dose level	Epirubicin (mg/m^2, day 1 q21d)	Capecitabine (mg/m^2 twice daily, days 1–14 q21d)	Incidence of DLTs*
Japanese CEX	4	100	900	—
	3	90	900	$\frac{2}{6}$
	2	90	829	$\frac{0}{3}$
	1	75	829	$\frac{1}{4}$
	0	75	628	$\frac{0}{3}$
EORTC CEX	4	100	1250	$\frac{2}{2}$
	3		1050	$\frac{0}{3}$
	2		900	$\frac{0}{3}$
	1		750	$\frac{0}{3}$

*The number of patients experiencing any DLT/the number of evaluable patients.

that the recommended CEX regimen be limited to dose level 2 and consist of capecitabine 900 mg/m^2 twice daily, epirubicin 100 mg/m^2, and CEX 600 mg/m^2.

Although the EORTC study identified a recommended CEX regimen in this way, concern was raised over possible differences in CEX tolerability between Caucasians and Japanese [6, 13]. To answer this question, we conducted a phase I dose-finding trial using the CRM to determine the RDs of the CEX combination in Japanese patients with advanced breast cancer [14, 15]. Based on data from the EORTC CEX study and assuming that the RD of CEX in Japanese patients should not be higher than that in Caucasians, five dose levels (0–4) were planned in the Japanese CEX study, as summarized in Table I. Treatment consisted of a fixed dose of CEX (600 mg/m^2 on day 1) in combination with three doses of epirubicin and three doses of capecitabine. Dose level 4, the highest in our study, corresponded to the CEX RD as determined in the EORTC CEX analysis. The European and Japanese CEX studies employed the same DLT definitions.

3. The CRM in the Japanese cex trial

3.1. Study design using the CRM

3.1.1. Dose–toxicity model. In the CRM we used numerical dose levels X_j for $j=0,\ldots,4$, to reduce the dimension of the dose levels for the CEX treatment consisting of the three anti-cancer agents. The numerical values of X_j were specified using 'backward fitting' [16] as described below, instead of the actual dose levels for the CEX treatment in Table I. This dimension reduction allows a dose–toxicity model to suitably fit the pre-study estimates of the proportion of patients who would experience a DLT at the dose levels. The outcome variable is the indicator $Y_i=1$ if a patient i suffers a DLT, 0 if not. A one-parameter logistic regression model,

$$\pi(X_i,\beta)=\Pr(Y_i=1|X_i,\beta)=\frac{\exp(\beta_0+\beta_1 X_i)}{1+\exp(\beta_0+\beta_1 X_i)} \tag{1}$$

with the intercept b_0 fixed at 3 and a slope parameter b_1, is assumed. The likelihood for n patients is

$$f(\mathbf{Y}_n|\mathbf{X}_n,\beta)=\prod_{i=1}^{n}\pi(X_i,\beta)^{Y_i}\{1-\pi(X_i,\beta)\}^{1-Y_i}. \tag{2}$$

3.1.2. Setting up the CRM. Before starting the study, we conducted a preliminary study among participating clinical oncologists to obtain necessary reference information for implementing the CRM. We set up the CRM design using the following five steps:

(i) In step 1, we identified the target DLT probability as 0.33 and obtained the prior estimates of the proportion of patients who would experience a DLT at each dose level from 0 to 4 as 0.05, 0.10, 0.25, 0.40, and 0.60, respectively.

(ii) In step 2, we predetermined the model's intercept b_0 at 3, as discussed in Section 3.1.3.
(iii) In step 3, we specified a prior distribution function of the slope b_1. Letting $Ga(a,b)$ denote the gamma distribution with mean a/b and variance a/b^2, we assumed $Ga(a,b)$ for b_1 in order to constrain the slope b_1 to be positive and for computational convenience. This constraint implies an assumption that a higher dose level increases the probability of DLT.
(iv) In step 4, we specified numerical values of X_j for $j=0,\ldots,4$ using backward fitting as follows. We added a constraint $E(b_1)=1$ that corresponds to an equation $a=b$ in the gamma prior distribution to make the *a priori* dose–toxicity curve exactly reflect the prior estimate of DLT occurrence probabilities regardless of the degree of clinical uncertainty [17]. Under the dose–toxicity model with the slope b_1 fixed at 1, we computed each X_j to match $\Pr(Y=1|X_j,\beta_0=3,\beta_1=1)$ with the prior probability estimate of DLT occurrence at dose level j for $j=0,\ldots,4$. As a result, $\{X_0,X_1,X_2,X_3,X_4\}=\{-5.94,-5.20,-4.10,-3.41,-2.60\}$.
(v) In step 5, we specified the hyperparameters of the prior $p(b_1|a,b)$ as $a=b=5$. Details of this step are described in Section 4.

3.1.3. Specification of the intercept b_0. Under $a=b=5$ and $b_0=3$, the prior dose-toxicity curve with a 90 per cent credible interval is given in Figure 1(a). This prior dose-toxicity curve may reflect the oncologist's greater confidence in higher rather than lower dose levels. That is, taking into account that dose level 4 in the Japanese CEX study corresponds to the RD identified in the EORTC CEX study, $b_0=3$ may be a reasonable choice. In contrast, if we use a negative value for the intercept, i.e. $b_0=-5$, $\{X_0,X_1,X_2,X_3,X_4\}$ is computed as $\{2.06, 2.80, 3.90, 4.59, 5.41\}$ using backward fitting. In this setting, the prior dose–toxicity curve represents greater uncertainty in higher rather than lower dose levels (Figure 1(b)) and therefore should be considered that the specification $b_0=-5$ contradicts the pre-study information.

3.1.4. Dose escalation/de-escalation rule. Our study plan involved treating up to 22 patients. The starting dose was level 1, which was given to the first enrolled patient. The CRM then ran sequentially with three patients per cohort. Each cohort was treated at the dose level X_j with an estimated probability of DLT $\pi\{X_j,E(\beta_1|\text{data})\}$ closest to 0.33 and not exceeding 0.40. If the computed probability of the suggested dose level was greater than 0.40, the cohort was treated at the preceding dose level. Untried doses were not skipped when escalating dose level. The trial was stopped if level 0 was considered too toxic to be administered, e.g. $\pi\{X_0,E(\beta_1|\text{data})\}>0.40$. The posterior distribution of the slope parameter b_1 and each posterior estimate $\pi\{X_j,E(\beta_1|\text{data})\}$ along with its 90 per cent credible

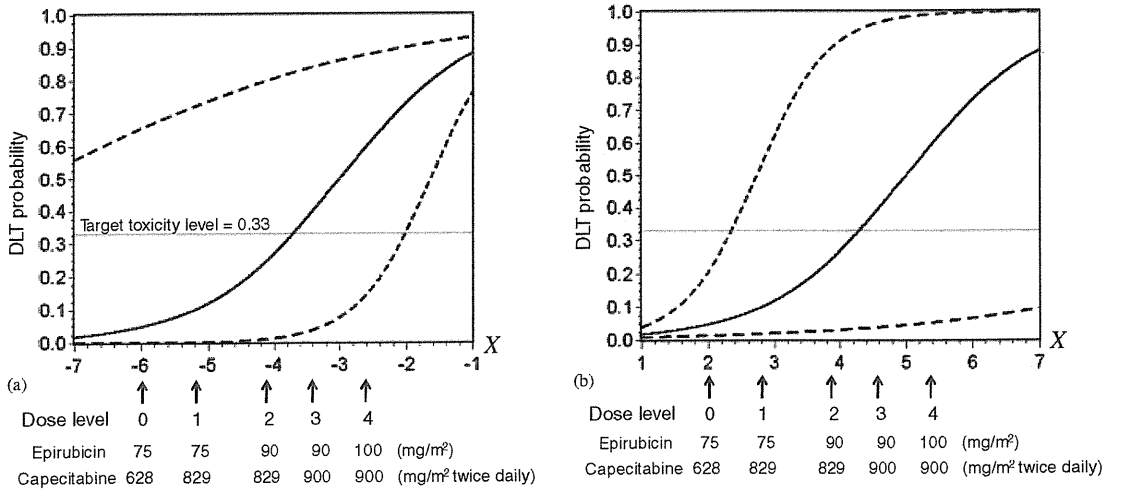

Figure 1. (a) Prior dose–toxicity curve (solid line) and its 90 per cent credible intervals (dashed lines) with the intercept $b_0=3$ under the gamma prior distribution, Ga (5,5). The horizontal axis X denotes the dose levels. The five values of $\{X1,X2,X3,X4,X5\}=\{-5.94,-5.20,-4.10,-3.41,-2.60\}$ used in the CRM computation are indicated by arrows. *The actual dose levels of epirubicin and capecitabine are also shown.* The horizontal straight line indicates the target DLT level (0.33) and (b) Prior dose–toxicity curve and its 90 per cent credible intervals with the intercept $b_0=-3$.

Table II. Incidence of dose-limiting toxicities (DLTs).

	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6
No. of evaluable patients	1	3	3	3	3	3
Doselevel*	1	0	1	2	3	3
Epirubicin (mg/m^2, day 1 q21d)	75	75	75	90	90	90
Capecitabine (mg/m^2 twice daily, days 1–14 q21d)	829	628	829	829	900	900
No. of patients experiencing any DLT	1	0	0	0	1	1
Grade 3 HFS†	1	—	—	—	—	—
Grade 3 anorexia	—	—	—	—	1	—
Grade 3 mucositis	—	—	—	—	—	1

*The dose level of CEX was fixed at 600 mg/m^2 on day 1 every 3 weeks.
†HFS, hand-foot syndrome.

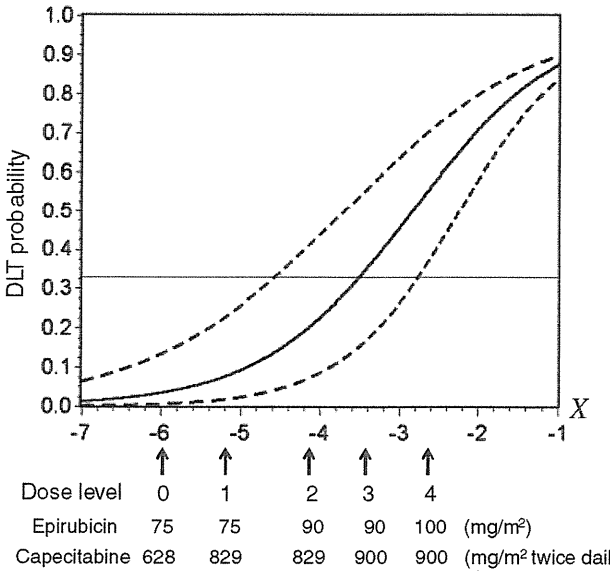

Figure 2. The posterior mean dose–toxicity curve (solid line) and its 90 per cent credible intervals (dashed lines) after updating with the toxicity data from all 16 patients.

interval were computed using numerical integration. An Independent Data and Safety Monitoring Committee (IDSMC) reviewed the interim analyses and was assigned the responsibility of making any recommendations to stop the trial on both clinical and statistical perspectives.

3.2. Implementation of the CRM

Because the results of the Japanese CEX trial were reported in detail in Saji *et al.* [14] and Morita *et al.* [15], we report here in brief. DLTs observed at each dose level and the dose escalation/de-escalation history throughout the study are shown in Tables I and II, respectively. The first patient treated at level 1 experienced a DLT (grade 3 hand-foot syndrome). The dose level was then de-escalated to level 0 for the second cohort. No DLTs were identified in the second, third (level 1), and fourth (level 2) cohorts. One of three patients in cohort 5 treated at level 3 experienced DLT (grade 3 anorexia). In the next cohort treated at level 3, one patient experienced DLT (grade 3 mucositis). Figure 2 shows the updated dose–toxicity curve including toxicity data from these 16 patients. The estimated DLT occurrence probability at level 3 was 0.354 (90 per cent credible interval: 0.174–0.560). With respect to efficacy data, one complete response and three partial responses were observed in six patients at level 3. Taking these CRM computations and the encouraging efficacy data into account, the DSMC recommended that the study be stopped. Therefore, we terminated the study and recommended that dose level 3 be further evaluated in a phase II trial.

4. Establishing a prior

In clinical trials with Bayesian model-based study designs, the prior should reasonably represent the physician's uncertainty. We established the prior distribution used in the Japanese CEX study based on the knowledge and experience of the participating clinical oncologists with regards to the CEX regimen. As described in Section 3, we assumed a gamma distribution Ga(a, b) for the prior distribution of the slope parameter b_1. Subject to $a=b$, the hyperparameter a determines the credible interval of the prior dose–toxicity curve under the gamma prior Ga(a, b). Thus, we determined that the hyperparameter a appropriately depicted the pre-study perceptions of the surveyed oncologists regarding the dose–toxicity relationship. By adjusting the hyperparameter a, i.e. $a=2$, 8, 20, 40, in addition to $a=5$ (Figure 1(a)) we created several graphical presentation patterns as shown in Figure 3. The clinical oncologists consulted in this study came to the consensus that the DLT probability at dose level 1 would be unlikely to be higher than 0.7 (more than double the target DLT level of 0.33) and the DLT probability at dose level 4 would be at least higher than 0.15 (around half of the target DLT level). The oncologists also concurred that the prior dose–toxicity curve and its credible interval constructed at $a=5$ reasonably reflected their knowledge and contained a sufficiently large degree of clinical uncertainty.

Although we determined the hyperparameters of the prior of b_1 based on an extensive discussion of the previous data using meticulous graphical presentations, our choice of the hyperparameters was arbitrary. If an established prior is overly informative, the prior may unduly influence posterior inferences and decisions, particularly early in the trial. Since dose levels must be selected sequentially in phase I dose-finding trials based on very small amounts of data, it may be important to quantify information contained in the chosen priors. These concerns may be addressed by quantifying the prior information

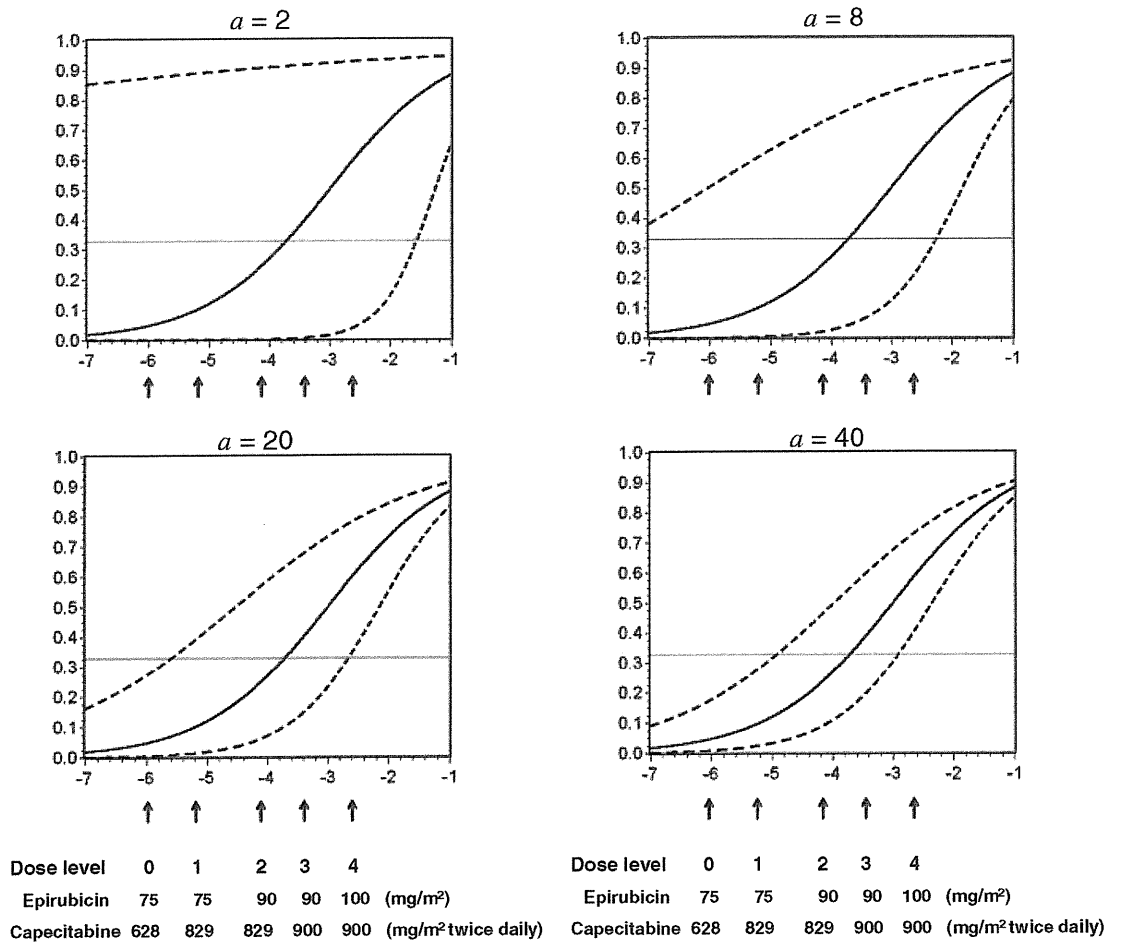

Figure 3. *Prior dose–toxicity curves with hyperparameters a=2, 8, 20, and 40.* Dashed lines indicate its 90 per cent credible intervals.

in terms of an equivalent number of hypothetical patients, i.e. a prior ESS. Such a summary would allow one to judge the relative contributions of the prior and the data to the decisions. We applied an ESS method proposed recently by Morita *et al.* [5] to the Japanese CEX trial in a retrospective fashion. The prior ESS computed at $a=5$ was 2.1. Thus, after enrolling three patients, the information from the likelihood started to dominate the prior, as desired. In addition, under Ga(5,5), the coefficient of variation (=standard deviation/mean) of the slope parameter b_1 was approximately 0.45, which might indicate some uncertainty in the slope parameter. Hence the prior specified in the Japanese CEX trial seemed quite reasonable.

As for the sensitivity analysis of the prior, the prior ESS values computed at $a=2$, 6, 7, 8, 20, and 40 are 0.86, 2.6, 3.0, 3.4, 8.6, and 17.1, respectively. It appears that $a<7$ may be needed to ensure an ESS<3. The prior with $a=40$ has ESS=17.1, so that it has impact roughly equal to that of the data on the posterior inference, as suggested by comparing Figures 2 and 3. In addition, under $a=40$, the *a priori* 90 per cent credible interval for the increase in the odds of a DLT occurrence, e.g. for the dose escalation from level 1 to level 2, is computed as 2.3–4.1, which may be excessively narrow compared with the 90 per cent credible interval of 1.5–7.5 computed under $a=5$. Thus, given the limited amount of information available during the design stage of the Japanese CEX study, the prior with $a=40$ may be criticized as being overly informative.

5. Discussion

When designing a phase I dose-finding study using a Bayesian CRM, certain choices must be made regarding details involved in a dose–toxicity model, numerical values of dose levels, prior distributions of model parameters, etc., and these should be sensible and plausible. If a one-parameter logistic model is chosen for modeling a dose–toxicity relationship, as was our approach in the Japanese CEX study, the intercept has to be specified at a certain real value. The actual dose levels of the combination therapy planned in the Japanese CEX study were based on information from the identical regimen conducted earlier in Caucasian patients, the EORTC CEX trial. In order to reduce the dimension of the dose levels, we specified the numerical values of the dose levels in the dose–toxicity formulation using backward fitting. In addition, we established the prior distribution of the slope parameter in the Japanese phase I trial by eliciting pre-study perceptions regarding the dose–toxicity relationship from Japanese clinical investigators.

So far, in many cases Japanese clinical investigators have conducted phase I studies assuming that a RD in Japanese patients should be lower than in Caucasian patients, based on results of clinical trials conducted in Western countries. That is, a large amount of historical data based on numerous studies has been integrated to design Japanese phase I trials. The Japanese CEX study, however, did not take full advantage of the pre-study information on dose–toxicity relationships derived from the EORTC CEX study to formally establish the prior distribution of the model parameter in the CRM.

Differences in RDs may be caused by specific differences between the abilities of Japanese and Caucasian populations to tolerate particular toxicities. These interracial differences can be regarded as patient prognostic covariates, but unfortunately such covariates have not yet been identified. Extensions of methods to find RDs for ordered prognostic subgroups have been proposed by O'Quigley and Paoletti [18], Yuan and Chappell [19], and Ivanova and Wang [20]. These methods may be applied to identifying RDs within racial subgroups in the setting of a multinational phase I study. Thall *et al.* [21] have proposed a Bayesian sequential phase I/II dose-finding design accounting for patient covariates and dose–covariate interactions. This method may also prove useful in modeling the Japanese–Caucasian association in a multinational study setting. It may be a significant challenge, however, to construct informative prior(s) on such an interracial difference in dose–toxicity curves [22].

In the context of Bayesian clinical trial design, well-chosen priors are important to ensure that posterior-based decision rules have good study operating characteristics. Some appropriate criteria for calibrating priors may be desired to obtain sensible prior distributions. A prior ESS quantifying the prior information in terms of the number of hypothetical patients may provide a useful tool for understanding the impact of prior-related assumptions. A useful property of prior ESS is that it is readily interpretable by clinical investigators who are involved in designing a clinical trial. ESS_RegressionCalculator.R, a computer program used to calculate the ESS for a normal linear or logistic regression model, is available from the website http://biostatistics.mdanderson.org/SoftwareDownload.

Acknowledgements

Satoshi Morita's research was supported in part by Grant H20-CLINRES-G-009 from the Ministry of Health, Labour and Welfare of Japan. He is grateful to Dr Masakazu Toi and Dr Junichi Sakamoto for providing important advice and suggestions. The author thanks the associate editor and the two referees for their constructive and helpful comments and suggestions.

References

1. Extra JM, Rousseau F, Bruno R, Clavel M, Le Bail N, Marty M. Phase I and pharmacokinetic study of taxotere (RP 56976; NSC 628503) given as a short intravenous infusion. *Cancer Research* 1993; **53**:1037–1042.
2. Taguchi T, Furue H, Niitani H, Ishitani K, Kanamaru R, Hasegawa K, Ariyoshi Y, Noda K, Furuse K, Fukuoka M, Yakushiji M, Kashimura M. Phase I clinical trial of RP 56976 (docetaxel) a new anticancer drug. *Japanese Journal of Cancer and Chemotherapy* (*Gan To Kagaku Ryoho*) 1994; **21**:1997–2005 (in Japanese).
3. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics* 1990; **46**:33–48.
4. O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics* 1996; **52**:673–684.
5. Morita S, Thall PF, Müller P. Determining the effective sample size of a parametric prior. *Biometrics* 2008; **64**:595–602.
6. Bonnefoi H, Biganzoli L, Cufer T, Mauriac L, Hamilton A, Schaefer P, Piccart M. An EORTC phase I study of epirubicin in combination with fixed doses of cyclophosphamide and infusional 5-fu (CEF-infu) as primary treatment of large operable or locally advanced/inflammatory breast cancer. *Breast Cancer Research and Treatment* 2001; **70**:55–63.
7. Bonnefoi H, Biganzoli L, Mauriac L, Cufer T, Schaefer P, Atalay G, Piccart M. An EORTC phase I study of capecitabine (Xeloda) in combination with fixed doses of cyclophosphamide and epirubicin (CEX) as primary treatment for large operable or locally advanced/inflammatory breast cancer. *European Journal of Cancer* 2003; **39**:277–283.
8. Blum JL, Jones SE, Buzdar AU, Lo Russo PM, Kuter I, Vogel C, Osterwalder B, Burger HU, Brown CS, Griffin T. Multicenter phase II study of capecitabine in paclitaxel-refractory metastatic breast cancer. *Journal of Clinical Oncology* 1999; **17**:485–493.
9. Blum JL, Dieras V, Lo Russo PM, Horton J, Rutman O, Buzdar A, Osterwalder B. Multicenter, phase II study of capecitabine in taxane-pretreated metastatic breast carcinoma patients. *Cancer* 2001; **92**:1759–1768.
10. O'Shaughnessy J, Miles D, Vukelja S, Moiseyenko V, Ayoub JP, Cervantes G, Fumoleau P, Jones S, Lui WY, Mauriac L, Twelves C, Van Hazel G, Verma S, Leonard R. Superior survival with capecitabine plus docetaxel combination therapy in anthracycline-pretreated patients with advanced breast cancer: phase III trial results. *Journal of Clinical Oncology* 2002; **20**:2812–2823.
11. Blum JL, Dees EC, Chacko A, Doane L, Ethirajan S, Hopkins J, McMahon R, Merten S, Negron A, Neubauer M, Ilegbodu D, Boehm KA, Asmar L, O'Shaughnessy JA. Phase II trial of capecitabine and weekly paclitaxel as first-line therapy for metastatic breast cancer. *Journal of Clinical Oncology* 2006; **24**:4384–4390.
12. Gradishar WJ, Meza LA, Amin B, Samid D, Hill T, Chen M. Capecitabine plus paclitaxel as front-line combination therapy for metastatic breast cancer: a multicenter phase II study. *Journal of Clinical Oncology* 2004; **22**:2321–2327.
13. Iwata H, Nakamura S, Toi M, Shin E, Masuda N, Ohno S, Takatsuka Y, Hisamatsu K, Yamazaki K, Kusama M, Kaise H, Sato Y, Kuroi K, Akiyama F, Tsuda H, Kurosumi M. Interim analysis of a phase II trial of cyclophosphamide, epirubicin and 5-fluorouracil (CEF) followed by docetaxel as preoperative chemotherapy for early stage breast carcinoma. *Breast Cancer* 2005; **12**:99–103.
14. Saji S, Toi M, Morita S, Iwata H, Ito Y, Ohno S, Kobayashi T, Hozumi Y, Sakamoto J. Dose-finding phase I and pharmacokinetic study of capecitabine (xeloda) in combination with epirubicin and cyclophosphamide (CEX) in patients with inoperable or metastatic breast cancer. *Oncology* 2007; **72**:330–337.
15. Morita S, Toi M, Saji S, Iwata H, Ohno S, Ito Y, Kobayashi T, Hozumi Y, Sakamoto J. Practical application of the continual reassessment method to a phase I dose-finding trial in advanced breast cancer. *Drug Information Journal* 2007; **41**:691–700.
16. Garrett-Mayer E. The continual reassessment method for dose-finding studies: a tutorial. *Clinical Trials* 2006; **3**:57–71.
17. Ishizuka N, Morita S. Practical implementation of the continual reassessment method. In *Handbook of Statistics in Clinical Oncology* (2nd edn), Crowley J (ed.). CRC Press: New York, 2005; 97–116.
18. O'Quigley J, Paoletti X. *Continual reassessment method for ordered groups. Biometrics* 2003; **59**:430–440.
19. Yuan Z, Chappell R. Isotonic designs for phase I cancer clinical trials with multiple risk groups. *Clinical Trials* 2004; **1**:499–508.
20. Ivanova A, Wang K. Bivariate isotonic design for dose-finding with ordered groups. *Statistics in Medicine* 2006; **25**:2018–2026.
21. Thall PF, Nguyen HQ, Estey EH. Patient-specific dose finding based on bivariate outcomes and covariates. *Biometrics* 2008; **64**:1126–1136.
22. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association* 2005; **100**:680–701.

Hepatology Research

Hepatology Research 2011; **41**: 296–302 doi: 10.1111/j.1872-034X.2011.00778.x

Original Article

Higher discontinuation and lower survival rates are likely in elderly Japanese patients with advanced hepatocellular carcinoma receiving sorafenib

Manabu Morimoto,[1] Kazushi Numata,[1] Masaaki Kondo,[1] Hisashi Hidaka,[2] Juichi Takada,[2] Akitaka Shibuya,[2] Satoshi Kobayashi,[3] Shinichi Ohkawa,[3] Chiaki Okuse,[4] Satoshi Morita,[5] Masataka Taguri[5] and Katsuaki Tanaka[1]

[1]Gastroenterological Center, Yokohama City University Medical Center, Yokohama, [2]Gastroenterology Division of Internal Medicine, Kitasato University East Hospital, Sagamihara, [3]Hepatobiliary and Pancreatic Medical Oncology, Kanagawa Cancer Center Hospital, Yokohama, [4]Gastroenterology and Hepatology, Department of Internal Medicine, St Marianna University School of Medicine, Kawasaki and [5]Department of Biostatistics and Epidemiology, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa, Japan

Aim: Sorafenib is approved for the treatment of advanced hepatocellular carcinoma (HCC) in Japan; however, its tolerability and efficacy in elderly patients with HCC have not been clarified. We aimed to evaluate the tolerability and efficacy of sorafenib with increasing age.

Methods: As part of a retrospective, multicenter cohort study conducted between May 2009 and February 2010, patients with advanced HCC received 400 mg sorafenib twice daily (standard dosage) or once daily (half-dosage) until disease progression or treatment intolerance.

Results: The mean age of the enrolled patients ($n = 76$) was 70.3 years, and 24 of them were ≥75 years old. The prognostic factors for survival were age < 75 years, performance status score zero, α-fetoprotein level < 1000 ng/mL, des-gamma-carboxy prothrombin level < 1000 ng/mL, and treatment duration ≥ 1 month. The median treatment duration and overall incidence of adverse drug reactions (ADRs) were not statistically different with increasing age. However, subgroup analysis revealed that treatment discontinuation because of ADRs was more frequent among the ≥75-year-old patients (41.7%) than among the <75-year-old ones (15.0%) with the standard dosage ($P = 0.047$); this trend was not observed among those who received the half-dose regimen.

Conclusions: Sorafenib has modest efficacy and acceptable toxicity in younger (<75 years) patients with HCC; however, elderly patients experience some side effects when it is administered at the standard dosage.

Key word: adverse drug reaction, dosage, elderly, hepatocellular carcinoma, sorafenib, survival

INTRODUCTION

HEPATOCELLULAR CARCINOMA (HCC) is the fifth most common type of cancer worldwide. It is highly prevalent in the Asia-Pacific region and Africa, and its incidence is increasing in Western countries.[1] Infection with hepatitis B or C virus is the greatest risk factor for hepatocarcinogenesis.

Sorafenib is the current standard drug for the first-line systemic treatment in patients with advanced HCC who are not candidates for curative treatments, such as surgical resection or locoregional therapies.[2] This multikinase inhibitor, with activity against Raf kinase and vascular endothelial cell growth factor (VEGF) receptor,[3] has been approved for the treatment of unresectable HCC by regulatory agencies of the European Union, United States, and other countries. This approval was based on the positive results of a placebo-controlled randomized phase III study of patients with advanced HCC.[4] Subsequently, a phase III study conducted in the Asia-Pacific region where hepatitis B virus infection is the predominant etiologic factor for chronic liver disease also demonstrated the survival benefits of sorafenib.[5]

Correspondence: Dr Manabu Morimoto, Gastroenterological Center, Yokohama City University Medical Center, 4-57 Urafune-cho, Minami-ku, Yokohama 232-0024, Japan. Email: morimoto@urahp.yokohama-cu.ac.jp
Received 10 October 2010; revision 6 December 2010; accepted 1 January 2011.

In global trials including non-Japanese populations, sorafenib was generally well tolerated;[4–7] however, their average age at presentation was relatively young (age range, 51–69 years). On the other hand, in our previous study of sorafenib treatment in Japanese patients with HCC,[8] the average age at presentation (70.3 years) was much older than in the previous trials,[4–7] and increasing age (≥75 years) was an important prognostic factor for lower overall survival (OS). At present, the efficacy and tolerability of this drug in elderly patients with advanced HCC is not clear; therefore, we conducted a secondary retrospective analysis of this multicenter trial[8] to evaluate the efficacy and tolerability of sorafenib with increasing age.

METHODS

Patients

THIS RETROSPECTIVE, MULTICENTER cohort study included patients with histopathologically and/or radiographically proven advanced HCC at four institutes of the Kanagawa Liver Study Group. All patients had measurable disease at baseline according to the response evaluation criteria in solid tumors (RECIST).[9] Further, all patients provided written informed consent. The institutional review board or ethics committee approved the study protocol, which complied with the Good Clinical Practice guidelines, the Declaration of Helsinki, and local laws.

Patients were excluded if they had previously received molecular-targeted therapies or any other systemic treatment. The inclusion criteria were Eastern Cooperative Oncology Group performance status (ECOG PS) score of 2 or less, Child–Pugh liver function class A or B, adequate hematologic function (platelet count > 5.0×10^{10}/L and hemoglobin level > 8.0 g/dL), adequate hepatic function (albumin level > 2.5 g/dL, total bilirubin level < 3.0 mg/dL, and alanine aminotransferase [ALT] and aspartate aminotransferase [AST] levels < five times the normal upper limit), and adequate renal function (serum creatinine level < 1.5 times the normal upper limit).

Treatment regimens

All of the patients received sorafenib between May 2009 and February 2010. The dosage was 400 mg twice daily (the standard dose); treatment interruptions and dose reductions (first 400 mg twice daily, then 400 mg once daily, and finally 400 mg every 2 days) were permitted for adverse drug reactions (ADRs). In some elderly patients (≥75 years) and those with poor liver function, the initial dose was reduced to half the standard dose, a 400-mg once-daily regimen. The patients received the therapy until any of the following criteria for discontinuation of therapy was met: ADRs that required termination of medication, disease progression, deterioration of ECOG PS score to 4, and withdrawal of consent. Other criteria for discontinuation included the concomitant use of an illicit drug that, in the opinion of the investigator, could induce toxicity or noncompliance with follow-up.

Response assessments

The patient response to treatment was evaluated according to the RECIST.[9] OS was measured from the date of administration of sorafenib until the date of death from any cause. The time to radiologic progression (TTRP) was defined as the time from the date of administration of sorafenib to disease progression, according to RECIST. Tumor measurements were performed at screening and every 4–6 weeks during treatment. Safety assessments included documentation of ADRs, clinical laboratory tests, physical examination, and measurement of vital signs. ADRs were defined according to Common Terminology Criteria for Adverse Events version 3.0 (CTCAE v3.0; http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf).

Statistical analysis

Continuous variables are represented as the mean ± standard deviation, and categorical variables are represented as the absolute and relative frequencies. The Mann–Whitney *U*-test was used to compare continuous variables between groups of patients; categorical variables were compared by using the Fisher's exact test or its equivalent for more than two categories. The TTRP and OS were calculated by Kaplan–Meier survival curves with log-rank survival comparisons and 95% confidence intervals (95% CI). Twenty-two variables were assessed using a univariate analysis to identify possible prognostic factors: age (≥70 years vs. <70 years and ≥75 years vs. <75 years), gender (male vs. female), etiology (hepatitis C vs. other), Child-Pugh class (A vs. B), tumor-node-metastasis (TNM) staging system revised by the Liver Cancer Study Group of Japan in 2008[10] (II or III vs. IV), tumor staging revised by Barcelona Clinic Liver Cancer (BCLC) group[11] (B vs. C), macrovascular invasion (absent vs. present), extrahepatic spread (absent vs. present), ECOG PS (score 0 vs. 1 to 2), initial dose of sorafenib (400 mg/day vs. 800 mg/day), total sorafenib

dose (≥30 000 mg vs. <30 000 mg), sorafenib-treatment duration (≥1 month vs. <1 month), average sorafenib dose (≥400 mg/day vs. <400 mg/day, ≥500 mg/day vs. <500 mg/day, and ≥600 mg/day vs. <600 mg/day), grade 3–4 ADRs (absent vs. present), platelet count (≥10 000 /μL vs. <10 000 /μL), serum albumin level (≥3.5 g/dL vs. <3.5 g/dL), α-fetoprotein (AFP) level (≥1000 ng/mL vs. <1000 ng/mL), des-gamma-carboxy prothrombin (DCP) level (≥1000 ng/mL vs. <1000 ng/mL), and treatment response according to the RECIST[9] (complete response, partial response, and stable disease vs. progressive disease). A Cox proportional hazards model was used to investigate prognostic factors for OS. All statistical analyses were carried out with the PASW Statistics 17.0 software (IBM SPSS, Inc., Chicago, IL, USA) and SAS (version 9.2).

RESULTS

Baseline characteristics

TABLE 1 SHOWS the baseline characteristics of the 76 patients enrolled in the study. Their mean age was 70.3 years (range, 37–88 years), and 24 (31.6%) patients were aged ≥75 years. Most of the patients (82.9%) were male, and 57 (75%) patients had a documented history of viral hepatitis (hepatitis B, hepatitis C, or both hepatitis B and C). Forty-one (53.9%) patients underwent transcatheter arterial chemoembolization, 10 (13.2%) received arterial infusion therapy, 13 (17.1%) received percutaneous ablation therapy, six (7.9%) underwent surgical resection, three (3.9%) received radiotherapy or other therapy, and three (3.9%) patients were never treated previously. Seventy-one (93.4%) patients presented with Child–Pugh class A liver cirrhosis and the remaining (6.6%) presented with Child–Pugh class B disease. Further, 24 (31.6%) patients had vascular invasion and 19 (25%) had extrahepatic spread of HCC. There were no significant differences in the baseline characteristics between the <75-year-old and the ≥75-year-old patients except in the case of previous therapy: arterial infusion chemotherapy was employed more often in the elderly patients ($P = 0.041$).

The standard dosage of sorafenib was administered to 52 (68.4%) patients and the half-dose regimen was administered to the remaining (31.6%). The patients in the latter group were significantly older and had Child-Pugh class B liver disease more often.

Safety and tolerability

Table 2 shows the incidence of ADRs (based on CTCAE v3.0) in relation to age and the treatment regimens. The incidence of any grade of ADRs (96.2% vs. 100%, $P = 0.230$) and grade 3–4 ADRs (44.2% vs. 54.2%, $P = 0.420$) were not significantly different between the <75-year-old and the ≥75-year-old patients. However, in the subgroup analysis of the two treatment regimens, anorexia (any grade) was significantly more common in the ≥75-year-old patients who received the standard dosage (75.0% vs. 22.5%, $P = 0.001$) compared with the half-dose regimen (50.0% vs. 16.7%, $P = 0.083$). Similarly, grade 3–4 anorexia was significantly more common in the ≥75-year-old patients who received the standard dosage (33.3% vs. 2.5%, $P = 0.001$) compared with the half-dose regimen (16.7% vs. 8.3%, $P = 0.537$).

The median treatment duration of all the patients was 1.7 months, without a significant difference between the age groups (1.9 months for <75 years vs. 1.4 months for ≥75 years). Sorafenib treatment was discontinued because of radiologic tumor progression and ADRs in 22 (42.3%) and nine (17.3%) of the <75-year-old patients, respectively, and eight (33.3%) and eight (33.3%) of the ≥75-year-old patients, respectively (Table 3). There was no statistical difference in the incidences of discontinuation due to radiologic tumor progression and ADRs between the <75-year-old and the ≥75-year-old patients ($P = 0.457$ and $P = 0.119$, respectively). In the subgroup analysis of the two treatment regimens, a higher percentage of the ≥75-year-old (41.7%) patients who received the standard dosage discontinued the therapy because of ADRs (vs. 15.0% of those <75 years old, $P = 0.047$); however, this trend was not observed in the half-dose regimen (25.0% for ≥75 years vs. 25.0% for <75 years, $P = 1.000$).

Efficacy and response

Overall, two (2.6%) patients had a complete response, three (3.9%) had a partial response, and 20 (26.3%) had stable disease. The response rate, defined as the percentage of patients with a complete or partial response, was 6.6%. Twenty-four deaths had occurred on endpoint during observation periods. The median OS of all the patients was 8.1 months (95% CI, 5.4–10.7), and the median TTRP was 2.9 months (95% CI, 2.0–3.7). A univariate analysis with a Kaplan–Meier model identified 10 variables as prognostic indicators of OS: age (<75 years vs. ≥75 years, $P = 0.022$), TNM staging system by the Liver Cancer Study Group of Japan (II+III vs. IV, $P = 0.027$), tumor staging by the BCLC group (B vs. C, $P = 0.015$), macrovascular invasion (absent vs. present, $P = 0.005$), ECOG PS (score 0 vs. 1–2, $P < 0.001$), total dose of sorafenib (≥30 000 mg vs. <30 000 mg, $P = 0.001$), treatment duration (≥1 month

Table 1 Demographic and baseline characteristics of the patients

Characteristics	<75-year-olds ($n = 52$)	≥75-year-olds ($n = 24$)	P
Age (years)			
Gender			0.945
Male	43 (82.7)	20 (83.3)	
Female	9 (17.3)	4 (16.7)	
Etiology			0.059
Hepatitis C only	28 (53.8)	16 (66.7)	
Hepatitis B only	10 (19.2)	0 (0.0)	
Hepatitis B and C	3 (5.8)	0 (0.0)	
Other	11 (21.2)	8 (33.3)	
Previous therapy			0.041
TACE	30 (57.7)	11 (45.8)	
Arterial infusion	3 (5.8)	7 (29.2)	
Percutaneous ablation	9 (17.3)	4 (16.7)	
Surgical resection	6 (11.5)	0 (0.0)	
Radiotherapy or others	3 (5.8)	0 (0.0)	
None	1 (1.9)	2 (8.3)	
ECOG PS score			0.110
0	38 (73.1)	14 (58.3)	
1	11 (21.2)	10 (41.7)	
2	3 (5.8)	0 (0.0)	
TNM stage			0.219
II	5 (9.6)	0 (0.0)	
III	21 (40.4)	13 (54.2)	
IVA	11 (21.2)	7 (29.2)	
IVB	15 (28.8)	4 (16.7)	
BCLC stage			0.876
B (intermediate)	25 (48.1)	12 (50.0)	
C (advanced)	27 (51.9)	12 (50.0)	
Child–Pugh class			0.564
A	48 (92.3)	23 (95.8)	
B	4 (7.7)	1 (4.2)	
Macrovascular invasion			0.451
Absent	37 (71.2)	15 (62.5)	
Present	15 (28.8)	9 (37.5)	
Extrahepatic spread			0.254
Absent	37 (71.2)	20 (83.3)	
Present	15 (28.8)	4 (16.7)	
Bone	7 (13.5)	1 (4.2)	
Lung	6 (11.5)	1 (4.2)	
Lymph nodes	1 (1.9)	1 (4.2)	
Other	1 (1.9)	1 (4.2)	
Biochemical analysis			
ALT (IU/L)	60 ± 54	45 ± 19	0.190
Total bilirubin (mg/dL)	1.0 ± 0.5	1.1 ± 0.5	0.686
Albumin (g/dL)	3.6 ± 0.5	3.5 ± 0.4	0.188
Platelets ($\times 10^4/\mu L$)	14.2 ± 7.2	15.9 ± 7.4	0.338
AFP (ng/mL)	13 791 ± 50 308	53 445 ± 142 974	0.198
AFP-L3 (%)†	33 ± 28	36 ± 27	0.805
DCP (ng/mL)‡	14 378 ± 54 696	13 108 ± 24 667	0.914

†$n = 48$.
‡$n = 75$.
The data represent the mean ± standard deviation or the number of patients (percentage).
AFP, α-fetoprotein; AFP-L3, fucosylated fraction of AFP; ALT, alanine aminotransferase; BCLC, Barcelona Clinic Liver Cancer staging system;[11] DCP, des-gamma-carboxy prothrombin; ECOG PS, Eastern Cooperative Oncology Group performance status; TACE, transcatheter arterial chemoembolization; TNM, tumor-node-metastasis staging revised by the Liver Cancer Study Group of Japan in 2008.[10]

Table 2 Incidence of adverse drug reactions (ADRs) according to age and treatment regimens

ADR	<75-year-olds			≥75-year-olds		
	Total	400 mg q.d. ($n = 12$)	400 mg b.i.d. ($n = 40$)	Total	400 mg q.d. ($n = 12$)	400 mg b.i.d. ($n = 12$)
All ADRs	96.2/44.2	100/50.0	95.0/42.5	100/54.2	100/41.7	100/66.7
Fatigue	63.5/3.8	41.7/8.3	70.0/2.5	50.0/8.3	58.3/16.7	41.7/0.0
Anorexia	21.2/3.8	16.7/8.3	22.5*/2.5*	62.5/25.0	50.0/16.7	75.0*/33.3*
Diarrhea	65.4/1.9	25.0/0.0	77.5/2.5	45.8/0.0	16.7/0.0	75.0/0.0
Hand-foot-skin reaction	53.8/13.5	41.7/16.7	57.5/12.5	33.3/0.0	33.3/0.0	33.3/0.0
Rash	13.5/0.0	25.0/0.0	10.0/0.0	8.3/0.0	8.3/0.0	8.3/0.0
Hypertension	15.4/0.0	25.0/0.0	12.5/0.0	25.0/0.0	16.7/0.0	33.3/0.0
ALT elevation	40.4/7.7	41.7/8.3	40.0/7.5	37.5/8.3	41.7/8.3	33.3/8.3
Bilirubin elevation	21.2/5.8	16.7/0.0	22.5/7.5	25.0/.12.5	16.7/0.0	33.3/25.0
Decreased platelet count	21.2/7.7	33.3/16.7	17.5/5.0	29.2/4.2	33.3/0.0	25.0/8.3

*Significant difference ($P = 0.001$) between the ≥75-year-old and the <75-year-old patients in the 400 mg b.i.d. regimen (standard dosage).
The data represent any grade (%)/grade 3–4 (%) of ADRs.
ALT, alanine aminotransferase.

vs. <1 month, $P < 0.001$), AFP level and DCP levels (<1000 ng/mL vs. ≥1000 ng/mL, $P < 0.001$), and treatment response (complete response, partial response, and stable disease vs. progressive disease, $P < 0.001$). A multivariate analysis with a Cox proportional-hazards model identified five variables as prognostic factors for OS: age (<75 years vs. ≥75 years; hazard ratio [HR], 0.237; 95% CI, 0.072–0.784; $P = 0.018$), ECOG PS (score 0 vs. 1–2; HR, 4.090; 95% CI, 1.113–15.037; $P = 0.034$), AFP level (<1000 ng/mL vs. ≥1000 ng/mL; HR, 0.131; 95% CI, 0.044–0.390; $P < 0.001$), DCP level (<1000 ng/mL vs. ≥1000 ng/mL; HR, 0.166; 95% CI, 0.047–0.578; $P = 0.005$), and treatment duration (≥1 month vs. <1 month; HR, 4.412; 95% CI, 1.016–19.159; $P = 0.048$). Figure 1 shows Kaplan–Meier curves of OS for <75-year-old and ≥75 year-old patients.

There were no significant differences in the average OS and the median TTRP between the patients receiving the standard-dose regimen and those receiving the half dose regimen (6.6 months vs. 5.8 months, $P = 0.965$ for OS; 3.0 months and 2.8 months, $P = 0.600$ for TTRP, respectively). In the subgroup analysis of patients ≥75 years, the average OS and the median TTRP were comparable between the two dose regimens (5.3 months vs. 5.0 months, $P = 0.839$ for OS; 2.0 months vs. 2.8 months, $P = 0.138$ for TTRP, respectively).

DISCUSSION

OUR PREVIOUS REPORT[8] indicated that median treatment duration and incidence of ADRs were not statistically different with increasing age; however, age ≥75 years was an important prognostic factor for lower OS. To reevaluate the relationship between patient age and drug safety, we conducted a secondary analysis using the same cohort and found that sorafenib has modest

Table 3 Incidence of treatment discontinuation according to age and treatment regimens

Reasons for discontinuation	<75-year-olds			≥75-year-olds		
	Total	400 mg q.d. ($n = 12$)	400 mg b.i.d. ($n = 40$)	Total	400 mg q.d. ($n = 12$)	400 mg b.i.d. ($n = 12$)
ADRs	9 (17.3)	3 (25.0)	6 (15.0)*	8 (33.3)	3 (25.0)	5 (41.7)*
Radiologic tumor progression	22 (42.3)	4 (33.3)	18 (45.0)	8 (33.3)	4 (33.3)	4 (33.3)

*Significant difference ($P = 0.047$) between the <75-year-old and the ≥75-year-old patients in the 400 mg b.i.d. regimen (standard dosage).
The data represent the number of patients (percentage).
ADR, adverse drug reaction.

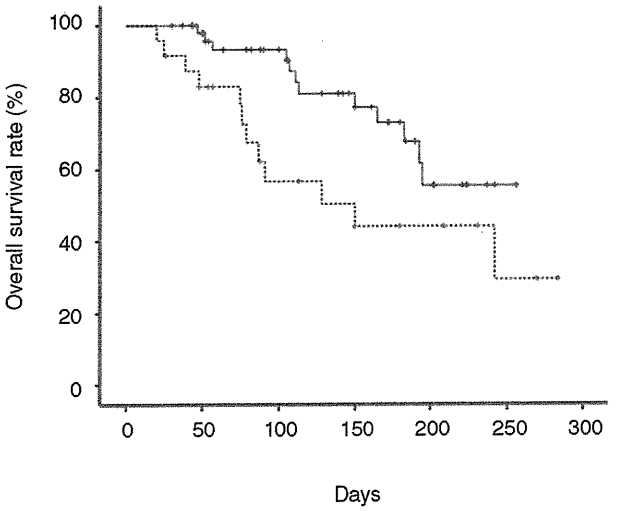

Figure 1 Kaplan–Meier estimates of overall survivals of the <75-year-old (solid line) and ≥75 year-old (broken line) patients. Univariate analysis revealed a significant difference ($P = 0.022$) between these age groups.

efficacy and tolerable ADRs in younger (<75 years) Japanese patients with advanced HCC; however, more than 40% of the elderly patients (≥75 years) who received the standard dosage (400 mg twice daily) discontinued the treatment because of ADRs. This is the first report indicating that older age is associated with a greater likelihood of discontinuation of sorafenib treatment and lower survival rate clinically.

The Raf/MAP kinase-ERK kinase (MEK)/extracellular signal-regulated kinase (ERK) pathway is overexpressed in HCC.[12–14] Sorafenib is a small molecule that inhibits multiple tyrosine kinases including Raf kinase, platelet-derived growth factor (PDGF), VEGF receptor 2 and 3 kinases, and c-Kit receptor, and it uniquely targets the Raf/MEK/ERK pathway.[15] Sorafenib was generally well tolerated in global trials of non-Japanese younger populations[4–7] and a phase I trial of Japanese patients.[16] In the present study, we demonstrated a median OS of 8.1 months and a median TTRP of 2.6 months. These data are similar to those of an Asia-Pacific trial[5] but do not indicate the benefit of sorafenib reported in the SHARP trial.[4] These conflicting results may be derived from the poorer treatment compliance in the current study than in the SHARP trial: 76% of the sorafenib-group patients received more than 80% of the planned daily dose of sorafenib in the SHARP trial, whereas only 40% of our patients received more than 80% of the planned daily dose (data not shown). Our study population also had significantly short treatment duration (median, 1.7 months) due to ADRs. The percentage of patients with any ADRs in this study was >10% higher than that in the global trials.[4,5] Especially, higher percentages of patients had fatigue, anorexia, and diarrhea (59.2%, 34.2%, and 59.2%, respectively) than those in the SHARP trial (22%, 14%, and 39%, respectively) or Asia-Pacific trial (20.1%, 12.8%, and 25.5%, respectively). These differences in ADRs can be explained by the differences in the elderly populations of these three studies: the mean age in the current study (70.3 years) is older than those in the SHARP trial (64.9 years) and Asia-Pacific study (51.0 years). On the other hand, our data indicate that the incidence of sorafenib discontinuation because of ADRs is low in younger patients (<75 years) (15.4%) and comparable with the published data.[4]

There are several reports indicating that the elderly are at increased risk of ADRs when they receive various antiangiogenic drugs.[17,18] van der Veldt *et al.*[17] showed that age and gender are predictive factors for severe toxicity of sunitinib in patients with advanced renal cell cancer. Ramalingam *et al.*[18] showed that, in elderly patients with non-small-cell lung cancer, the addition of bevacizumab to standard chemotherapy did not improve the clinical outcome but results in increased toxicity and treatment-related deaths compared with patients aged <70 years. Therefore, we emphasize that when taking the antiangiogenic drugs, including sorafenib, for adjuvant or maintenance treatment in elderly patients with HCC, the special concern of the safety might be needed.

In the present study, the difference in the initial dose did not affect the OS and TTRP. A recent case report of a 74-year-male patient with advanced HCC described that the patient received the half-dose of sorafenib for 8 months and achieved a more than 16 months survival benefit without disease recurrence.[19] Therefore, further studies will be needed to determine whether a reduced-dose regimen of sorafenib truly imparts a survival benefit for patients with HCC comparable to the standard-dose regimen. At present, older age alone should not preclude the therapeutic option using standard-dose of sorafenib; however, such regimen might be cautious for elderly patients with other risk factors to avoid discontinuation of sorafenib due to ADRs.

The Cox proportional-hazards model indicated age (<75 years), ECOG PS (score 0), AFP level (<1000 ng/mL), DCP level (<1000 ng/mL), and treatment duration (≥1 month) as favorable prognostic factors for OS. In a global phase III trial of patients with renal cell carcinoma treated with sorafenib, multivariate analysis