



Published in final edited form as:

Cancer. 2011 June 15; 117(12): . doi:10.1002/cncr.25750.

## Progression-Free Survival as a Predictor of Overall Survival in Metastatic Renal Cell Carcinoma Treated With Contemporary Targeted Therapy

Daniel Y. C. Heng, MD, MPH<sup>1</sup>, Wanling Xie, MS<sup>2</sup>, Georg A. Bjarnason, MD<sup>3</sup>, Ulka Vaishampayan, MD<sup>4</sup>, Min-Han Tan, MD<sup>5</sup>, Jennifer Knox, MD<sup>6</sup>, Frede Donskov, MD<sup>7</sup>, Lori Wood, MD, MPH<sup>8</sup>, Christian Kollmannsberger, MD<sup>9</sup>, Brian I. Rini, MD<sup>10</sup>, and Toni K. Choueiri, MD, MSc<sup>2,11</sup>

<sup>1</sup>Tom Baker Cancer Center, University of Calgary, Calgary, Alberta, Canada

<sup>2</sup>Dana Farber Cancer Institute, Boston Massachusetts

<sup>3</sup>Sunnybrook Odette Cancer Center, Toronto, Ontario, Canada

<sup>4</sup>Karmanos Cancer Institute, Wayne State University, Detroit, Michigan

<sup>5</sup>National Cancer Center, Singapore

<sup>6</sup>Princess Margaret Hospital, University of Toronto, Toronto, Ontario, Canada

<sup>7</sup>Aarhus University Hospital, Aarhus, Denmark

<sup>8</sup>Dalhousie University, Halifax, Nova Scotia, Canada

<sup>9</sup>British Columbia Cancer Agency, Vancouver, British Columbia, Canada

<sup>10</sup>Cleveland Clinic Taussig Cancer Institute, Cleveland, Ohio

<sup>11</sup>Department of Medical Oncology, Dana Farber Cancer Institute, Boston, Massachusetts

### Abstract

**BACKGROUND**—The majority of metastatic renal cell carcinoma (mRCC) clinical trials that examined targeted agents used progression-free survival (PFS) as the primary endpoint. Whether PFS can be used as a predictor of overall survival (OS) is unknown.

**METHODS**—Patients from 12 cancer centers who received targeted therapy for mRCC were identified. Landmark analyses for progression at 3 months and 6 months after drug initiation were performed to minimize lead-time bias. A proportional hazards model was used to assess the utility of PFS for predicting OS.

**RESULTS**—In total, 1158 patients were included. The median follow-up was 30.6 months, the median age was 60 years, and the median Karnofsky performance status was 80%. For the entire

---

© 2010 American Cancer Society.

**Corresponding author:** Daniel Y. C. Heng, MD, MPH, Tom Baker Cancer Center, Alberta Health Services Cancer Care, University of Calgary, 1331 29th Street NW, Calgary, Alberta, T2N 4N2 Canada; Fax: (403) 283-1651; daniel.heng@albertahealthservices.ca.

#### CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

This is a project of the International Metastatic Renal Cell Carcinoma (mRCC) Database Consortium.

D.Y.C.H., W.X., B.I.R., and T.K.C. were involved in the conception, study design, statistical analysis, and writing of the article.

D.Y.C.H., G.A.B., U.V., M.-H.T., J.K., F.D., L.W., C.K., B.I.R., and T.K.C. provided and collected patient data. All authors approved the final article.

cohort, the median PFS was 7.6 months, and the median OS was 19.7 months. In the landmark analysis, the median OS for patients who progressed at 3 months was 7.8 months compared with 23.6 months for patients who did not progress (log-rank test;  $P < .0001$ ). Similarly, for patients who progressed at 6 months, the median OS was 8.6 months compared with 26 months for patients who did not progress ( $P < .0001$ ). Compared with those who did not progress, for the patients who progressed at 3 months and at 6 months, the hazard ratios for death adjusted for adverse prognostic factors were 3.05 (95% confidence interval, 2.42-3.84) and 2.96 (95% confidence interval, 2.39-3.67), respectively. Similar results were demonstrated with landmark analyses at 9 months and at 12 months and in the bootstrap validation. Kendall tau rank correlation and a Fleischer model demonstrated a statistically significant dependent correlation.

**CONCLUSIONS**—PFS at 3 months and at 6 months predicted OS, and the current results indicated that PFS may be a meaningful intermediate endpoint for OS in patients with mRCC who receive treatment with novel agents.

### Keywords

progression-free survival; overall survival; surrogate endpoints; vascular endothelial growth factor; metastatic renal cell carcinoma

---

**Metastatic** renal cell carcinoma (mRCC) has been revolutionized by the introduction of agents that target the vascular endothelial growth factor (VEGF) and mammalian target of rapamycin (mTOR) pathways. Progression-free survival (PFS) has been used as the primary endpoint for the majority of clinical trials that investigate these patients. Trials involving sunitinib, sorafenib, bevacizumab, everolimus, and pazopanib demonstrated a PFS benefit, which subsequently led to their regulatory approval.<sup>1-6</sup> With the exception of 1 trial that involved temsirolimus in patients with a poor prognosis,<sup>7</sup> the vast majority of these trials did not prove an overall survival (OS) benefit and cited issues of crossover and contamination of the control arm. This potentially was because patients in the control arm were allowed to receive the active agent or were treated with similar second-line and third-line agents, which may have led to the dilution of any apparent OS benefit. Because of the emergence of second-line and third-line targeted therapies, an OS benefit may be difficult to prove in a trial setting.

In the current study, we sought to determine whether there is an association between PFS and OS, which is the traditional gold standard of all endpoints, and whether or not there is any dependence between these 2 endpoints. This has implications for future drug evaluation and trial design.

## MATERIALS AND METHODS

### Study Population

In total, 1158 patients with mRCC who received treatment with contemporary targeted therapy (sunitinib, sorafenib, bevacizumab, temsirolimus) were included in this study. They were identified from consecutive, population-based patient samples during 2005 through 2009 at 12 international cancer centers in Canada (Toronto Sunnybrook Odette Cancer Center [n = 158], Princess Margaret Hospital [n = 95], British Columbia Cancer Agency [n = 100], Queen Elizabeth II Health Sciences Center [n = 76], and Alberta Health Services Cancer Care [n = 134]), the United States (Dana Farber Cancer Institute and Brigham and Women's Hospital [n = 177], Cleveland Clinic [n = 113], Karmanos Cancer Center [n = 89], Beth Israel Deaconess Medical Center [n = 74]), Singapore (National Cancer Center [n = 80]), and Denmark (Aarhus University Hospital [n = 62]).

Patients may have been treated on clinical trial or off protocol and may have been treated at major academic centers or community oncology centers. Baseline patient characteristics and outcome data were collected using uniform data-collection templates in this large retrospective analysis.<sup>8</sup> Regulatory approval from local institutional review boards or research ethics boards was obtained for each center.

## Endpoints

The primary endpoint was OS, which was defined as the time from drug initiation to the date of death from any cause or was censored at last the follow-up. PFS was defined as the time from drug initiation to the time of disease progression or death or was censored at last the follow-up. We defined disease progression according to standard Response Evaluation Criteria in Solid Tumors (RECIST).<sup>9</sup>

## Data Analysis

Landmark analyses of PFS at 3 months and 6 months after drug initiation were performed to minimize lead-time bias. These time points were defined a priori because they were likely to coincide with the routine radiographic assessment of patients who receive VEGF-targeted therapy. Thus, this would be clinically useful to determine whether disease progression at these early time points would be able to predict a difference in OS. Studies in other tumors, such as prostate cancer, have justified the use of 3-month and 6-month landmark analyses.<sup>10</sup> The patients who died before 3 months or 6 months were excluded in the 3-month landmark analysis or the 6-month landmark analysis, respectively, to provide the most conservative estimate of effect.<sup>10</sup>

The method of Kaplan and Meier was used to estimate the OS of patients stratified by disease progression at 3 months or 6 months. Exploratory analyses also were performed at 9 months and at 12 months, but those analyses were subjected to lead-time bias. A proportional hazards model was used to assess the significance of progression at 3 months or 6 months in predicting OS when adjusted for poor prognostic factors, including a Karnofsky performance status (KPS) <80%, time from diagnosis to treatment <1 year, anemia, hypercalcemia, thrombocytosis, and neutrophilia.<sup>8</sup> Similar analyses were performed that adjusted for Memorial Sloan-Kettering Cancer Center (MSKCC) poor prognostic criteria.<sup>11</sup> The proportionality assumption was met by graphically assessing plots of log(-log[survival]) versus log of survival time. The case deletion method was used to handle missing values in all explanatory variables.

The correlation between PFS and OS was estimated using the statistical model for dependence between PFS and OS developed by Fleischer et al.<sup>12</sup> That model follows a parametric approach, which investigates how much variability of OS can be explained by variability from PFS. In the model, OS is partitioned as a sum of the time to progression (TTP ~ Exp[ $\lambda_1$ ]) and postprogression survival (PPS ~ Exp[ $\lambda_3$ ]) if patients progress before dying. In patients who have not experienced progression, PFS is equal to OS (Exp[ $\lambda_2$ ]). The correlation between PFS and OS is obtained by:

$$\text{corr}(PFS, OS) = \frac{\lambda_3}{\sqrt{\lambda_1^2 + 2\lambda_1\lambda_2 + \lambda_3^2}} = \sqrt{\frac{\text{Var}(PFS)}{\text{Var}(OS)}}$$

where model parameters  $\lambda$  can be estimated using a plug-in method for the estimated medians. Because this model requires an exponential distribution for time variables, we graphically assessed this assumption using the plot of negative log of survival distribution

(LS) against time. This curve should be approximately linear through the origin if the exponential model is appropriate.<sup>12</sup>

We also investigated the nonparametric Kendall tau rank correlation for bivariate censored data,<sup>13</sup> which has been used in previous studies of surrogate endpoints in prostate cancer.<sup>10</sup> This is a more conservative measure and does not require data conforming to certain distributions. However, the Kendall tau considers PFS and OS as independent events and ignores the reality that OS inherently contains PFS information. The Kendall Tau correlation was calculated by using the `cenken` function from the R NADA library (<http://cran.r-project.org/web/packages/NADA/index.html>) modified for right double censored time to event data.

The standard errors of the correlation statistics (Fleischer model and Kendall tau) were estimated using the bootstrap method with 300 replications. The bootstrap confidence intervals (CIs) were computed based on the percentiles of the bootstrap distribution of the statistic. For the landmark analysis, 300 bootstrap samples were generated similarly at each landmark time point. We refit the Cox regression models and calculated the adjusted hazard ratios as in the original analysis. The bootstrap means and CIs were computed and compared with the model using the original study population at each landmark time point.

Bootstrap analysis was done using the statistical software package R version 2.8.0 (<http://cran.r-project.org>). All other statistical analyses were undertaken using SAS version 9.2 (SAS Institute Inc., Cary, NC). All statistical tests were 2-sided, and *P* values <.05 were considered significant.

## RESULTS

### Baseline Characteristics

In total, 1158 patients were included in the current analysis. At the time of the analyses, the median follow-up of the entire cohort was 30.6 months, and 81% of patients had progressed. The median KPS was 80% and the median age was 60 years. Thirty-one percent of patients had received previous immunotherapy, 80% had previously undergone nephrectomy, 77% had >1 metastatic site, and 8% had brain metastases. The risk groups according to the criteria published by Heng et al<sup>8</sup> were favorable, intermediate, and poor for 21.1%, 50.2%, and 28.7% of patients, respectively. MSKCC risk groups<sup>11</sup> were favorable, intermediate, and poor in 21.2%, 56.2% and 22.6% of patients, respectively (Table 1).

### Landmark Analyses

In the landmark analysis, the median OS for patients who progressed at 3 months was 7.8 months compared with 23.6 months for the patients who did not progress (log-rank test; *P* < .0001) (Fig. 1). Similarly, for the patients who progressed at 6 months versus those who did not progress, the median OS was 8.6 versus 26.0 months, respectively (*P* < .0001). For the patients who progressed at 3 months and 6 months, compared with those who did not progress, the hazard ratios for death adjusted for adverse prognostic factors<sup>8</sup> were 3.05 (95% CI, 2.42-3.84) and 2.96 (95% CI, 2.39-3.67), respectively. Hazard ratios when adjusted for the MSKCC prognostic groups (data not shown) and additional landmark analyses at 9 and 12 months (Table 2) all revealed consistent findings. Bootstrap analyses that were performed at 3 months, 6 months, 9 months, and 12 months produced very similar hazard ratios, confirming the high internal validity of these results. All of the aforementioned hazard ratios had a *P* value <.0001.

## Correlation Analyses

For all patients ( $n = 1158$ ), the median PFS was 7.6 months (95% CI, 6.8-8.2 months), and the median OS was 19.7 months (95% CI, 18.1-21.6 months). The median post-progression survival (PPS) for patients who progressed before death was 8.5 months (95% CI, 7.8-9.3 months). The negative log survival distribution curves followed a straight line through the origin, indicating satisfaction of the exponential distribution (data not shown). By using the Fleischer model, the estimated correlation between PFS and OS was 0.66 (bootstrap standard error [SE], 0.025; 95% CI, 0.61-0.71). This means that approximately 44% ( $r^2 = 0.66^2$ ) of the variability in OS can be explained by PFS. The calculated Kendall tau correlation was 0.42 (bootstrap SE, 0.016; 95% CI, 0.39-0.45;  $P < .0001$ ), which suggested a statistically significant dependence between PFS and OS.

## DISCUSSION

In the landmark analyses at each time point, we demonstrated that patients with mRCC who progressed on contemporary targeted therapy had an approximately 3 times increased risk of dying compared with patients at the same time point who remained progression free. A significant positive relation between PFS and OS was observed, as demonstrated by the Kendall tau statistic (0.42) and the Fleischer model (0.66). The Kendall tau statistic is a nonparametric measure of association between 2 censored variables and likely underestimated the true correlation between PFS and OS, because it ignored the inherent reality that OS contains PFS. The Fleischer model was developed specifically to describe this relation and, thus, may be a better estimate.<sup>12</sup>

To determine whether our results were consistent with expected values, we used a recent simulation study by Broglio and Berry,<sup>14</sup> which suggested that the correlation between PFS and OS relies on the duration of postprogression survival (PPS). When the median PPS is short, the hazard ratios for progression and OS are highly correlated. According to simulation results, the estimated correlation was 0.88 when the median PPS was 9 months, which is higher than what we observed in our data (0.66). However, the Broglio and Berry<sup>14</sup> simulation correlation was from data that perfectly followed an exponential distribution analyzed in a simulated meta-analysis format instead of using individual PFS and OS data for each patient as was done in our study.

The current analysis is important, because the use of OS as the gold-standard endpoint has become difficult in contemporary clinical trials. Many patients with mRCC in the control arms of randomized controlled trials eventually crossover to the active treatment or get contaminated by second-line or third-line drugs in the same drug category. This contamination can be as high as 62%, as documented in studies that examined bevacizumab plus interferon<sup>4</sup> and sunitinib.<sup>15</sup> There may be ethical concerns of not allowing crossover of patients on the control arm of clinical trials at the first hint of clinical activity on the experimental arm. However, if patients do crossover, then the OS benefit may be diluted. Thus, PFS may be the only endpoint that is not affected by contamination and crossover.

Intermediate endpoints like PFS and disease-free survival have been validated in other cancers, such as adjuvant and metastatic treatments for colorectal cancer,<sup>16-18</sup> and similar analyses have been published for prostate cancer.<sup>10</sup> Although most regulatory agencies and public health systems prefer to use OS data to make important approval decisions, PFS may be a suitable endpoint if it can be correlated with OS. This is especially true if there is a large magnitude of difference in PFS between the 2 treatment arms. It is important to validate these surrogate endpoints individually in each disease.

The strength of our study is that, to our knowledge, it represents the largest collection of patients that includes only those who received targeted therapy. Although 35.2% of patients in our study in fact did receive active second-line and/or third-line therapies, we still were able to demonstrate a correlation between PFS and OS using our study design. Most important, there was consistently high internal validity, as demonstrated by bootstrap validation.

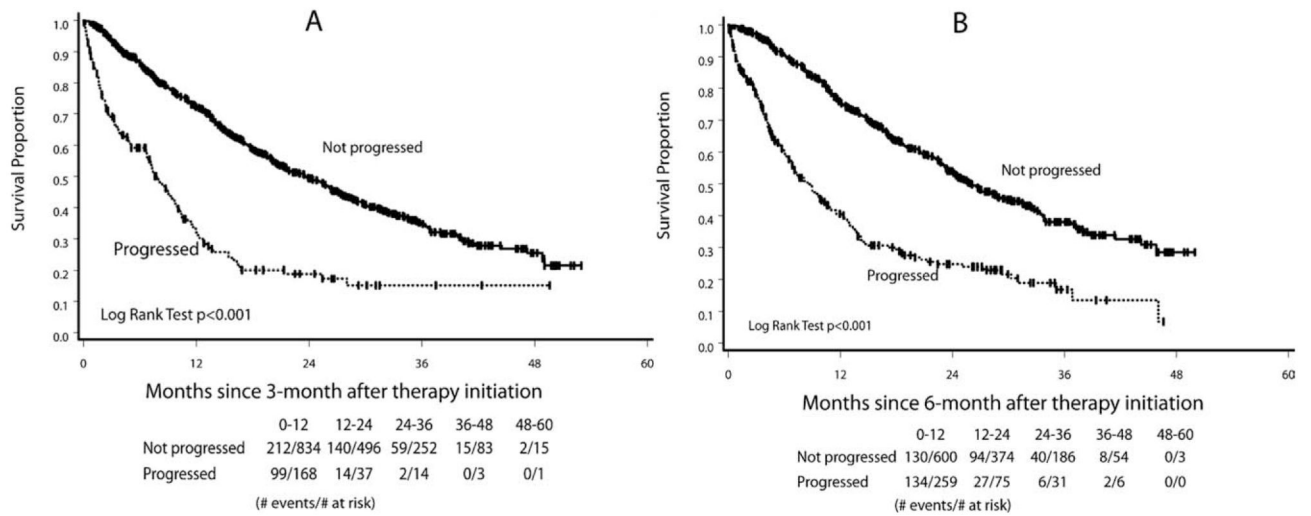
Limitations of this study include its retrospective nature, for which missing data and selection bias may be encountered. However, standardized data collection templates were used, and only 5 patients had missing PFS and/or OS data. In addition, these patients constituted a consecutive series of patients from each institution, thereby limiting selection bias. In fact, the inclusion of consecutive patients created a patient population that was diverse and that was not limited to clinical trials (approximately 15% were included in a randomized controlled trial, not including expanded access programs); thus, this may be a more generalizable study. Finally, there was no central radiology review to determine progression in this study. However, this actually may better reflect everyday physician practices in the routine management of metastatic RCC and, thus, would make the results more applicable.

We conclude that disease progression at 3 months and at 6 months is associated with and can independently predict OS and that there is a dependent relation between PFS and OS. PFS may be the only endpoint that is not hindered by the issues of contamination and crossover that were observed for OS in most trials of targeted therapy, in which subsequent second-line and third-line targeted therapy were commonplace. These results require prospective evaluation.

## REFERENCES

1. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med*. 2007; 356:115–124. [PubMed: 17215529]
2. Escudier B, Koralewski P, Pluzanska A, et al. A randomized, controlled, double-blind phase III study (AVOREN) of bevacizumab/interferon- $\alpha$ 2a vs placebo/interferon- $\alpha$ 2a as first-line therapy in metastatic renal cell carcinoma [abstract]. *J Clin Oncol*. 2007; 25(18S) Abstract 3.
3. Escudier BJ, Bellmunt J, Negrier S, et al. Final results of the phase III, randomized, double-blind AVOREN trial of first-line bevacizumab (BEV) + interferon- $\alpha$ 2a (IFN) in metastatic renal cell carcinoma (mRCC) [abstract]. *J Clin Oncol*. 2009; 27(15S) Abstract 5020.
4. Rini BI, Halabi S, Rosenberg JE, et al. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. *J Clin Oncol*. 2008; 26:5422–5428. [PubMed: 18936475]
5. Motzer RJ, Escudier B, Oudard S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet*. 2008; 372:449–456. [PubMed: 18653228]
6. Sternberg CN, Davis ID, Mardiak J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol*. 2010; 28:1061–1068. [PubMed: 20100962]
7. Hudes G, Carducci M, Tomczak P, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N Engl J Med*. 2007; 356:2271–2281. [PubMed: 17538086]
8. Heng DY, Xie W, Regan MM, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. *J Clin Oncol*. 2009; 27:5794–5799. [PubMed: 19826129]
9. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009; 45:228–247. [PubMed: 19097774]

10. Halabi S, Vogelzang NJ, Ou SS, et al. Progression-free survival as a predictor of overall survival in men with castrate-resistant prostate cancer. *J Clin Oncol.* 2009; 27:2766–2771. [PubMed: 19380448]
11. Motzer RJ, Bacik J, Murphy BA, et al. Interferon-alfa as a comparative treatment for clinical trials of new therapies against advanced renal cell carcinoma. *J Clin Oncol.* 2002; 20:289–296. [PubMed: 11773181]
12. Fleischer F, Gaschler-Markefski B, Bluhmki E. A statistical model for the dependence between progression-free survival and overall survival. *Stat Med.* 2009; 28:2669–2686. [PubMed: 19579225]
13. Akristas MG, Siebert J. A test for partial correlation with censored astronomical data. *Mon Not R Astron Soc.* 1995; 278:919–924.
14. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst.* 2009; 101:1642–1649. [PubMed: 19903805]
15. Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2009; 27:3584–3590. [PubMed: 19487381]
16. Sargent DJ, Wieand HS, Haller DG, et al. Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *J Clin Oncol.* 2005; 23:8664–8670. [PubMed: 16260700]
17. Tang PA, Bentzen SM, Chen EX, et al. Surrogate end points for median overall survival in metastatic colorectal cancer: Literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *J Clin Oncol.* 2007; 25:4562–4568. [PubMed: 17876010]
18. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol.* 2007; 25:5218–5224. [PubMed: 18024867]



**Figure 1.** These charts illustrate landmark analyses for overall survival according to progression status (A) at 3 months and (B) at 6 months after the initiation of targeted therapy.



**Table 1**Patient and Disease Characteristics at the Initiation of Targeted Therapy (N=1158)<sup>a</sup>

<b>Variable</b>	<b>Total Cohort (n=1158), %</b>
Median age [range], y	60 [13-89]
Median KPS [range]	80 [30-100]
Men	73.4
>1 Metastatic site	77.4
Brain metastases	8
Prior nephrectomy	80.2
Targeted therapy as first-line therapy	68.6
<b>Type of targeted therapy</b>	68.6
Sunitinib	69.3
Sorafenib	22.7
Bevacizumab	6.6
Temsirolimus	1.2
Axitinib	0.2
<b>MSKCC risk category</b>	
Favorable	21.2
Intermediate	56.2
Poor	22.6
<b>Heng criteria risk</b>	
Favorable	21.1
Intermediate	50.2
Poor	28.7

KPS indicates Karnofsky performance status; MSKCC, Memorial Sloan-Kettering Cancer Center.

<sup>a</sup>The risk category was unknown in 195 patients (MSKCC) and in 125 patients (Heng criteria; Heng 2009<sup>8</sup>).

**Table 2**

Results of Landmark and Bootstrap Analyses for Overall Survival by Progression Status at 3 Months, 6 Months, 9 Months, and 12 Months After the Initiation of Targeted Therapy (N=1158)

Landmark, Time, mo	No. of Patients Excluded <sup>b</sup>	No. Who Did Not Progress/ Progressed	Median OS After Landmark Time, mo <sup>a</sup>		Adjusted HR (95% CI) <sup>a</sup>	Bootstrap Mean (95% CI) for Adjusted HR
			Did Not Progress	Progressed		
3	156	834/168	23.6	7.8	3.05 (2.42-3.84)	3.15 (2.45-4.09)
6	299	600/259	26.0	8.6	2.96 (2.39-3.67)	3.06 (2.40-3.87)
9	425	445/288	28.7	9.7	2.83 (2.26-3.56)	2.91 (2.17-3.60)
12	542	330/286	32.0	11.0	3.24 (2.49-4.20)	3.29 (2.57-4.29)

OS indicates overall survival; HR, hazard ratio; CI, confidence interval.

<sup>a</sup> All OS times were compared in a log-rank analysis and revealed *P* values <.0001. All adjusted HRs that were calculated using the Heng criteria revealed *P* values <.0001.

<sup>b</sup> Reasons for exclusion: 1) death before landmark time (n=109 [3 mo], n=214 [6 mo], n=285 [9 mo], and n=362 [12 mo]), 2) follow-up for survival or disease not reaching landmark time (n=42 [3 mo], n=80 [6 mo], n=135 [9 mo], n=175 [12 mo]), and 3) missing OS/progression-free survival data (n=5).