Clinical Therapeutics/Volume 37, Number 11, 2015

An Indirect Comparison of Everolimus Versus Axitinib in US Patients With Advanced Renal Cell Carcinoma in Whom Prior Sunitinib Therapy Failed

Steven Sherman, MPH1; Billy Amzal, MSc, MPA, PhD1; Emiliano Calvo, MD, PhD2; Xufang Wang, MD, MBA3; Jinhee Park, PhD3; Zhimei Liu, PhD3; Chinjune Lin, MD3; and Roman Casciano, MSc1

1LASER Analytica, New York, New York; 2Centro Integral Oncológico Clara Campal and START Madrid, Madrid, Spain; and 3Novartis Pharmaceuticals Corporation, East Hanover, New Jersey

ABSTRACT

Purpose: The purpose of this study was to perform a weight-adjusted indirect comparison to approximate the relative efficacy of everolimus versus axitinib among patients with second-line metastatic renal cell carcinoma in whom sunitinib therapy previously failed.

Methods: Individual patient data from the RECORD-1 (Renal Cell Cancer Treatment With Oral RAD001 Given Daily) Phase III clinical trial provided information for patients taking everolimus. Summary baseline clinical and demographic characteristics and progression-free survival (PFS) outcomes were available for patients taking axitinib who were included in the AXIS (axitinib versus sorafenib) Phase III clinical trial. A Bayesian latent class mixture model differentiating responders and nonresponders and with imbedded Weibull regression on PFS was used to identify sex, Memorial Sloan-Kettering Cancer Center risk score, and time receiving prior sunitinib therapy as prognostic factors for PFS based on posterior probability >95%. Patients taking everolimus were weighted up or down based on their combination of prognostic variables. Weights were calculated by dividing the proportion of patients observed in AXIS for a given characteristic by the proportion observed in RECORD-1 and taking the product of the values derived for all three weighting variables considered. Weighted PFS distributions were derived with bootstrapped 95% CIs and compared with those reported for the AXIS trial.

Findings: After weighting, distributions of the 3 key baseline characteristics were more closely aligned between the 2 studies; however, some differences remained. A slightly lower rate of poor-risk patients was evident in RECORD-1 (30%) versus AXIS (36%), and a 9% lower proportion of males was observed in the everolimus group compared with the axitinib group. Distributions of time receiving prior sunitinib therapy were almost equivalent between the treatment arms. A median PFS of 4.7 months (95% CI, 3.5–10.6 months) was observed for patients in the weighted everolimus group compared with 4.8 months (95% CI, 4.5–6.4 months) in the AXIS trial.

Implications: Similar median PFS point estimates and overlapping CIs suggest that everolimus and axitinib have similar efficacy. Although these results do not negate the need for direct comparison, this study may be used to inform clinical and reimbursement decisions until such evidence is available. (Clin Ther. 2015;37:2552–2559) © 2015 The Authors. Published by Elsevier HS Journals, Inc.

Key words: axitinib, everolimus, indirect comparison, relative efficacy, renal cell carcinoma.

INTRODUCTION

Metastatic renal cell carcinoma (mRCC) is one of the leading causes of cancer deaths in the United States1; however, in recent years, several therapy options have become available to treat advanced RCC or mRCC. First-line treatments with National Comprehensive Cancer Network category 1 recommendations include sunitinib, temsirolimus, bevacizumab plus interferon alfa, pazopanib, high-dose interleukin 2, and sorafenib,2 although recent real-world evidence suggests that

Accepted for publication September 27, 2015.
http://dx.doi.org/10.1016/j.clinthera.2015.09.013
0149-2918/$ - see front matter
© 2015 The Authors. Published by Elsevier HS Journals, Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
sunitinib currently represents the dominant first-line treatment.\textsuperscript{3} Despite the observed progression-free survival (PFS) benefit from first-line sunitinib therapy, most patients require a second-line treatment within 1 year.\textsuperscript{4} After failure of first-line sunitinib therapy, according to National Comprehensive Cancer Network guidelines, everolimus and axitinib are the only agents that currently have category 1 evidence for second-line treatment of mRCC.\textsuperscript{2}

Everolimus, a mammalian target of rapamycin pathway inhibitor, was approved by the US Food and Drug Administration in 2009 as a second-line treatment for advanced RCC after failure of the vascular endothelial growth factor tyrosine kinase inhibitor sunitinib and sorafenib.\textsuperscript{5} In the pivotal Phase III RECORD-1 (Renal Cell Cancer Treatment With Oral RAD001 Given Daily) clinical trial, patients with mRCC in whom first-line sunitinib and/or sorafenib therapies previously failed were assigned to treatment with either everolimus or placebo plus best supportive care. The PFS and safety profiles were assessed for each treatment group until progression, death, or discontinuation, whichever came first. The study resulted in a significant PFS improvement of 67\% (hazard ratio [HR] = 0.33; 95\% CI, 0.25-0.43, \( P < 0.001 \)) for patients treated with everolimus versus placebo.\textsuperscript{6} A similar trend was observed in those with sunitinib as their only prior targeted therapy, with a median PFS of 4.6 versus 1.8 months (HR = 0.22; 95\% CI, 0.09-0.55) for patients treated with everolimus and best supportive care, respectively.\textsuperscript{7}

Axitinib is a vascular epidermal growth factor receptor 1 to 3, c-KIT, and platelet-derived growth factor receptor inhibitor approved in 2012 to treat advanced RCC after failure of prior systemic therapy.\textsuperscript{8} The AXIS (axitinib vs sorafenib) Phase III clinical trial examined the PFS and safety profiles among patients who had progressed with a first-line regimen of sunitinib, bevacizumab plus interferon alfa, temsirolimus, or cytokine therapy. Patients were randomized to receive either axitinib or sorafenib treatment. Overall, patients had significant improvement with axitinib compared with sorafenib, with a median PFS of 6.7 months for patients receiving axitinib versus 4.7 months for patients receiving sorafenib (HR = 0.665; 95\% CI, 0.544-0.812, \( P < 0.0001 \)).\textsuperscript{9} Similarly, a statistically significant PFS benefit was also found in favor of axitinib over sorafenib in the subgroup of patients previously treated with only sunitinib, with a median PFS of 4.8 months for axitinib and 3.4 months for sorafenib (HR = 0.741; 95\% CI, 0.573-0.958, \( P = 0.0107 \)).\textsuperscript{9}

Overall, both everolimus and axitinib therapies appear to provide PFS benefit for patients with advanced RCC who previously progressed with sunitinib treatment; however, because the overall study populations for the RECORD-1 and AXIS trials are not directly comparable due to population heterogeneity, a naive comparison of the trial results cannot be used to draw conclusions regarding relative efficacy in this context. This study aims to address this gap by performing a weight-adjusted indirect comparison to balance clinical and demographic characteristics between the 2 pivotal Phase III trial populations. Similar weighting methods have previously been used to balance patient characteristics among different studies when patient-level data are available for one treatment and only summary-level data for the comparator treatment.\textsuperscript{10,11} Although performing a randomized clinical trial is preferable, such studies can provide valuable insight until direct evidence becomes available.

**METHODS**

Data from the RECORD-1\textsuperscript{6} and the AXIS\textsuperscript{12} Phase III clinical trials were used to perform a weight-adjusted indirect comparison to align population characteristics and compare PFS for everolimus versus axitinib. Both clinical trials were multicenter, double-blind randomized clinical trials that evaluated PFS as their primary end point. Individual patient data were available from the RECORD-1 trial for everolimus patient follow-up, whereas only summary data were available for analytical purposes in this analysis for axitinib-treated patients. PFS estimates for the prior-sunitinib subgroup of the AXIS clinical trial (\( N = 194 \)) were obtained from the AXIS publication,\textsuperscript{9} whereas aggregated patient characteristics for those receiving axitinib were retrieved from data reported in a summary of the Evidence Review Group report on the manufacturer’s National Institute for Health Care Excellence submission.\textsuperscript{12} Survival data were extracted by digitizing the available Kaplan-Meier curve using the open source Engauge Digitizer software, version 4.1 (Engauge, Atlanta, Georgia). The overall everolimus population in the RECORD-1 trial (\( N = 277 \)) consisted of patients with potentially multiple prior lines of therapy. A subset (\( n = 43 \)) of patients with second-line mRCC receiving everolimus who were sunitinib refractory was identified in RECORD-1 to correspond with patients with second-line sunitinib-refractory mRCC receiving axitinib from the AXIS trial (\( n = 194 \)).

November 2015 2553
After limiting the RECORD-1 population to patients with second-line sunitinib-refractory mRCC, heterogeneity in the patient characteristics between the axitinib- and everolimus-treated patient groups persisted (Table I). Further adjustments were necessary to improve the comparability between the 2 studies; however, because patient characteristics and outcomes for patients receiving axitinib were reported only as summary measures, any attempt to align population characteristics between the everolimus and axitinib treatment arms was limited to adjustments to the RECORD-1 data.

In the data exploration, a bimodal distribution of survival times was observed, mixing predominantly 2 groups of responders (ie, slow vs fast progressors). As a consequence, a mixture of the 2 Weibull

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>RECORD-1 Prematched sunitinib-refractory second-line everolimus group (n = 43)</th>
<th>Postmatched sunitinib-refractory second-line everolimus group (n = 43)</th>
<th>AXIS sunitinib-refractory axitinib group (n = 194)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (IQR)</td>
<td>58 (32–79)</td>
<td>58 (32–79)</td>
<td>61 (20–82)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>70</td>
<td>65</td>
<td>74</td>
</tr>
<tr>
<td>Female</td>
<td>30</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td>Ethnic origin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>95</td>
<td>96</td>
<td>NR</td>
</tr>
<tr>
<td>Black</td>
<td>–</td>
<td>0</td>
<td>NR</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>2</td>
<td>NR</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>2</td>
<td>NR</td>
</tr>
<tr>
<td>ECOG performance status score†</td>
<td>0 42</td>
<td>1 58</td>
<td>&gt; 1 0</td>
</tr>
<tr>
<td>MSKCC risk subgroups‡</td>
<td>19 58</td>
<td>23 0</td>
<td>88 33</td>
</tr>
<tr>
<td>Prior nephrectomy</td>
<td>88</td>
<td>93</td>
<td>33</td>
</tr>
<tr>
<td>Prior radiation</td>
<td>33</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td>Time receiving prior sunitinib, mo</td>
<td>&lt; 9 79</td>
<td>&lt; 9 48</td>
<td>&lt; 9 50</td>
</tr>
</tbody>
</table>

AXIS = axitinib versus sorafenib; ECOG = Eastern Cooperative Oncology Group; IQR = interquartile range; MSKCC = Memorial Sloan-Kettering Cancer Center; NA = not applicable; NR = not reported; RECORD-1 = Renal Cell Cancer Treatment With Oral RAD001 Given Daily.

*Data are presented as percentage of patients unless otherwise indicated.
†RECORD-1 distribution based on Karnofsky Performance Status score < 80 as ECOG performance status score of 1.
‡Prognostic variables used in the weighting algorithm.
distributions describing the mix of these 2 subpopulations was evaluated. This approach allowed for the estimation of a latent classifying parameter for each patient together with their probability to be fast or slow progressor.

To account for the bimodal nature of the survival distribution using patient data for the true second-line sunitinib-refractory RECORD-1 population \((N = 43)\), a latent class mixture model was used with imbedded logit regression to reflect the mixture of responders and nonresponders. Latent class mixture models are useful for modeling heterogeneous longitudinal data by assigning patients into subgroupings.13

Conditionally to responder status, PFS was modeled as a Weibull regression with prognostic variables included as factors in a linear regression of the logarithm of the parameter scale defining the Weibull function. The following candidate covariates were considered: age, sex, Eastern Cooperative Oncology Group (ECOG) performance status, Memorial Sloan-Kettering Cancer Center (MSKCC) risk score, prior nephrectomy, prior radiation therapy, and time receiving prior sunitinib. The covariates included in the logistic regression to define responder status were those from the list above that had statistical significance \((P < 0.05)\) in the Weibull model without considering the latent class variable, namely, sex and prior nephrectomy.

The mixture latent class model was then fitted in a Bayesian set-up with noninformative uniform priors. On the basis of a posterior distribution on parameters \(> 95\%\), MSKCC risk score, sex, time receiving prior sunitinib, and prior nephrectomy were identified as influential variables for PFS outcomes. These variables were evaluated for independence using \(\chi^2\) tests of independence comparing all factors with each other. Through these exercises, MSKCC risk score, sex, and time receiving prior sunitinib were identified as key characteristics to weigh to balance the RECORD-1 patient data to match the summary characteristics reported in the AXIS publications. Prior nephrectomy was not considered as a weighting characteristic because of observed colinearity with the time receiving prior sunitinib variable.

Weights were calculated for each patient by dividing the proportion \((P_i)\) of patients observed in AXIS for a given key characteristic by the proportion observed in RECORD-1 and multiplying these quotients for all 3 weighting variables considered, according to the following equation:

\[
W_i = \frac{P_i(A_{AXIS})}{P_i(A_{RECORD-1})} \times \frac{P_i(B_{AXIS})}{P_i(B_{RECORD-1})} \times \frac{P_i(C_{AXIS})}{P_i(C_{RECORD-1})}
\]

where A indicates sex; B, MSKCC risk; and C, time receiving prior sunitinib. The weight option in the standard \textit{phreg} procedure using SAS statistical software, version 9.3 (SAS Institute Inc, Cary, North Carolina), was used to account for weighting and derive a new survival distribution for the adjusted everolimus population. This new weighted survival curve was then used to determine median PFS measures for patients with second-line disease receiving sunitinib and everolimus. This estimate was compared with the median PFS observed for patients receiving axitinib in the AXIS trial. The 95% CIs were determined through bootstrapping techniques using Matlab, release 14 (Mathworks, Natwick, Massachusetts). The latent class model was fitted with WinBUGS software, version 1.4.3 (MRC Biostatistics Unit, Cambridge, and Imperial College School of Medicine, London, England). All other aspects of this analysis were performed using SAS statistical software, version 9.3.

**RESULTS**

The distribution of derived study weights were right skewed with a mean (SD) of 1.0 (0.74) and a median (interquartile range) of 0.76 (0.5–1.1) (Figure 1). After applying weights, the distributions of the 3 key baseline characteristics were more closely aligned between the 2 studies; however, some differences remained (Table I). The RECORD-1 mean age remained slightly lower for the everolimus population compared with the axitinib population. Both the everolimus and axitinib populations had 23% of patients with a history of radiation therapy. A slightly lower rate of poor-risk patients based on MSKCC criteria was present in RECORD-1 (30%) versus AXIS (36%), and a 9% lower proportion of males was observed in the postweighting everolimus group compared with the axitinib population. Distributions of time receiving prior sunitinib therapy were almost equivalent between the everolimus and axitinib treatment arms.

A median PFS of 4.7 months (95% CI, 3.5–10.6 months) was estimated for the weighted everolimus...
patient group compared with 4.8 months (95% CI, 4.5–6.4) reported in the AXIS trial (Figure 2). Results of the indirect analysis suggest overlapping PFS CIs between everolimus and axitinib.

**DISCUSSION**

After identifying key drivers of PFS and adjusting for differences between the AXIS and RECORD-1 clinical trials, the results of the weighted-adjusted indirect comparison suggest similar PFS between

![Figure 1. Distribution of study weights.](image1)

![Figure 2. Kaplan-Meier progression-free survival curves for patients in the axitinib and weighted everolimus groups.](image2)
everolimus and axitinib among second-line patients with RCC treated previously with sunitinib. This analysis provides a more robust estimation of the relative efficacy between everolimus and axitinib compared with a naive comparison that ignores differences in trial inclusion criteria and distributions of prognostic variables. These findings are consistent with those reported in a recently published study.14

The indirect comparison analysis has several limitations. Randomized clinical trials remain the gold standard for clinical research, and as such, these results are only intended to approximate a clinical trial evaluating these 2 agents and should only be interpreted as those from an observational study. As with any observational study, bias may result from observed and unobserved confounding. The MSKCC risk score was defined differently in the RECORD-1 and AXIS trials. Although both trials used high corrected calcium levels and low hemoglobin levels as risk factors contributing to the overall risk score, the RECORD-1 study used a Karnofsky Performance Status (KPS) of <80 as an additional risk factor, whereas the AXIS trial considered ECOG performance status (PS) >0 (Table II). Levels of KPS do not directly correspond with levels of ECOG PS, and as such, it is not possible to convert between KPS to ECOG PS and maintain the same information (ie, an ECOG PS of 0 requires no symptoms and no restrictions, whereas a KPS of 80 requires symptoms and effort) (Figure 3). This difference in how MSKCC risk factors were determined likely led to inflated MSKCC risk levels in the AXIS study population, which may lead to an overestimation of PFS relative to everolimus. Furthermore, the limited number of second-line sunitinib-refractory patients in RECORD-1 (N = 43) likely limited the precision of the derived PFS estimate for the matched everolimus treatment group, leading to a wide CI of the estimate. The small sample size in the everolimus analysis population may also have adversely affected the ability to obtain exact distributional matches after weighting.

This analysis used a naive estimator to evaluate the probability of a patient being in a group given certain characteristics, which is highly dependent on sample size to perform well.15 Weighting variables were therefore limited to only those found to be prognostic for PFS so as not to compromise the functionality of this estimator. In doing so, prognostic variables identified by statistical modeling of RECORD-1 data were assumed to also be prognostic in the AXIS trial population. Although sex, MSKCC risk, and time receiving prior sunitinib were identified through this process, sample size limitations may have limited the ability to detect an association between other variables and PFS. In addition, the multivariate modeling revealed an association between PFS and time receiving prior sunitinib therapy; however, this result is not consistent with real-world findings where no association was found.16,17 This could be explained by the very different contexts in which the studies were performed (ie, it is possible that the results observed in this study may only hold true in the controlled environment of a randomized clinical trial).

This approach also required the assumption that residual differences in nonprognostic variables were negligible. This may not always be the case. For example, a higher proportion of Asians in the overall AXIS trial population versus RECORD-1 may have led to bias in the PFS result in favor of axitinib over everolimus because subanalyses in the AXIS trial suggest a stronger PFS treatment effect in nonwhites (HR = 0.524; 95% CI, 0.338–0.812) compared with whites (HR = 0.733; 95% CI, 0.475–1.138).

### Table II. Differences in MSKCC risk score calculation between the RECORD-1 and AXIS clinical trials.

<table>
<thead>
<tr>
<th>Risk Factors for the MSKCC Score used in RECORD-1</th>
<th>Risk Factors in the MSKCC Score used in AXIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected calcium</td>
<td>Corrected calcium</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>Hemoglobin</td>
</tr>
<tr>
<td>KPS score &lt;80</td>
<td>ECOG performance status score &gt;0</td>
</tr>
</tbody>
</table>

AXIS = axitinib versus sorafenib; ECOG = Eastern Cooperative Oncology Group; KPS = Karnofsky Performance Status; MSKCC = Memorial Sloan-Kettering Cancer Center; RECORD-1 = Renal Cell Cancer Treatment With Oral RAD001 Given Daily.

*A patient’s risk score was computed as the number of risk factors.
CI, 0.587–0.916). Additional head-to-head studies evaluating everolimus versus axitinib should be performed to confirm these results.

**CONCLUSIONS**

Similar median PFS point estimates (4.7 vs 4.8 months) with overlapping CIs suggest that everolimus and axitinib have similar efficacy. Although the study findings should not be considered conclusive without confirmatory direct head-to-head evidence, indirect comparison studies can be used to support clinical and reimbursement decisions, particularly in an ever-changing treatment landscape. Indirect comparison studies can be used to support clinical and reimbursement decisions, particularly in an ever-changing treatment landscape.

In this context, because the study results suggest similar efficacy in the population analyzed, reimbursement decisions may focus on overall cost of therapy and the potential effect of the safety profile. Because new data on these products are generated for the population with second-line sunitinib-refractory disease receiving everolimus, this analysis may be reproduced with a larger study population and more statistical power. Such studies may be used to further support any future reimbursement decisions and to elucidate the most appropriate and cost-effective sequence of therapy in this setting.

**ACKNOWLEDGMENTS**

We thank Allison Perrin and Andrew Chua for their assistance and coordination in the development of the manuscript.

**AUTHOR CONTRIBUTIONS**

Steven Sherman was responsible for study conceptualization and design, data analysis, data interpretation, drafting of manuscript. Billy Amzal was responsible for methodology review, Bayesian analysis, data interpretation, and review of manuscript drafts. Emiliano Calvo, Xufang Wang, Jinhee Park, Chinjune Lin, were responsible for study design, data interpretation, manuscript revision and approval. Zhimei Liu was responsible for data interpretation, manuscript revision and approval. Roman Casciano was responsible for study conceptualization and design, data interpretation, review of manuscript drafts.

**FUNDING SOURCE**

Novartis Pharmaceuticals Corporation provided funding for this study.

**CONFLICT OF INTEREST**

The sponsored research did not put limits on freedom to publish or the content of publication. S. Sherman, B. Amzal, and R. Casciano are
employees of LASER Analytica, a consultancy that received compensation for the overall economic study design, the analysis, and preparation of this manuscript.

X. Wang, J. Park, Z. Liu, C. Lin, are employees of and own stock in Novartis Pharmaceuticals Corporation. They also contributed to the analysis and manuscript preparation. EC is a consultant for Novartis Pharmaceuticals Corporation and has received compensation for his contributions to the analysis and the preparation of the manuscript. The authors have indicated that they have no other conflicts of interest regarding the content of this article.

REFERENCES


Address correspondence to: Roman Casciano, MSc, 875 Third Ave, Suite 6A, New York, NY 10022. E-mail: roman.casciano@la-ser.com