Reviewer's report

**Title:** Metastatic renal cell cancer treatments: An indirect comparison meta-analysis

**Version:** 2  **Date:** 3 November 2008

**Reviewer:** Paul Elson

**Reviewer's report:**

**Strengths:**

This is a very timely report on an important issue to physicians treating patients with advanced renal cell carcinoma. The authors conducted a thorough review of several databases to identify randomized trials of new anti-angiogenic agents in an effort to get some sense of their relative efficacy to each other. The study uses a contemporary set of five published studies and two abstracts to address the issue; and a reasonable statistical approach is used that takes advantage of the randomized nature of the trials but also allows for comparison between treatments even though none were conducted in a formal head-to-head fashion. The authors suggest that sunitinib is superior to sorafenib and bevacizumab but present conflicting results regarding sorafenib versus bevacizumab. They present their results as hypothesis-generating and conclude only that the new agents offer improved progression-free survival compared to interferon or placebo.

**Major Compulsory Revisions**

Most of the authors' results regarding the comparison of the new anti-angiogenic agents to each other are based on a single trial of each drug. In one case the trial is a randomized phase II study with less than 100 patients per treatment arm. This is too small a number of trials to really address differences between them that could account for at least some of the observed effects; and the authors should emphasize that the study is hypothesis generating and not meant to be conclusive. For example, the phase II trial referred to above is the only trial of sorafenib versus interferon included in the analysis. Randomized phase II trials are inherently different in terms of their goals from phase III trials and it is not clear that the results from one should necessarily be compared to the results of the other. Also there are differences in the way the primary outcome, progression-free survival, was defined. Most trials defined it as time from randomization to documented progression or death; however for at least one trial it appears to be based solely on whether or not the patient progressed (Yang, 2003). Similarly although the authors indicate that they used investigator assessments of progression-free survival, for at least one study the hazard ratio reported appears to be the one based on the assessments conducted by an independent panel (Escudier (sorafenib), 2007). Also, while most of the summary statistics the authors used for their analysis are based on simple comparisons of new agent versus interferon/placebo, in one instance an adjusted estimate was
used (Rini, 2008). Finally, there are differences between the studies with respect to the distribution of a very strong prognostic factor, risk group status. For example, in the Rini (2008) study of bevacizumab 26% of patients were considered favorable risk by the Memorial Sloan-Kettering criteria and 10% were unfavorable, whereas 36% of patients in the Motzer (2007) trial were favorable risk and 7% were unfavorable. This issue is relevant because the statistical method the authors used to make indirect comparisons is based in part on the assumption that there is no interaction between the magnitude of the treatment effect and the predictor. It is not clear, however that this is a safe assumption in this setting.

The data in Tables 1 and 2, (and possibly Figures 2 and 4) need to be reviewed. For example:

a) The patient population for the 2007 Escudier trial of sorafenib is described as being restricted to intermediate and poor risk patients; however the manuscript indicates that low and intermediate risk patients were eligible.

b) The study by Yang is described as having been conducted in 2003; however the reference gives the date as 2007.

c) The study by Hudes is described as being restricted to patients with performance status scores of 60-70. According to the manuscript patients had to have scores >60.

d) In Table 2 the column labeled “RR” only reports objective response rates for one study; for most others odds ratios appear to be presented.

e) Is the hazard ratio (1.14; interferon favored over sorafenib) for the Szczylik (2007) study correct? The confidence interval is not consistent with this, but is consistent with the inverse (0.88; sorafenib favored over interferon). If the latter is correct this could impact a number of the comparisons (and Figures 2 and 4).

The discussion of the temsirolimus trial and its inclusion in Tables 1 and 2 and Figure 2 should perhaps be dropped. As the authors point out it was conducted in a different patient population than the other trials; and there is no comparison of temsirolimus to any of the other agents. In addition, from Table 1 the authors appear to have combined the results from the temsirolimus and temsirolimus+interferon arms; and it is not clear that this is reasonable since it in theory could lead to no effect or an over (under) estimation of the impact of the drug compared to interferon depending on whether the effect of interferon in the combination is negligible, positive, or negative. If it is left in the authors should clarify the statement on page 13 that the study was restricted to intermediate and high risk patients. While this is correct based on Memorial Sloan-Kettering risk group definitions it appears to contradict the description of the study in Table 1 which indicates that the trial was restricted to “modified poor risk” patients. Similarly page 14 indicates that the patients were all previously treated whereas Table 1 indicates the trial was restricted to untreated patients.

In the first paragraph of the “Data Analysis” section the authors indicate that they
used hazard ratios as reported in the relevant publications, based on communication with the authors (was it always the lead or senior author?), or based on their own recalculation. This is an important aspect of the study since not all the data can be found in the papers/abstracts; and the authors should indicate in Table 2 (perhaps using footnotes) what the sources were for the different statistics.

In that same paragraph the authors indicate that they combined studies when more than one trial of a particular agent was available. Bevacizumab is the only agent for which this occurred and only two trials were involved. Given the indirect comparison method the authors used it is not immediately apparent why the two trials should be combined for the purpose of comparing the effects of bevacizumab to those of sunitinib and sorafenib.

The authors indicate on page 14 that Tables 1 and 2 suggest that duration of treatment, PFS, and objective response were all similar between patient groups. However the statement is not really supported by the data. For example treatment duration is missing for three of the six non-temsirolimus trials; and Table 2 does not report response rates per se for most of the trials; making comparison difficult.

In the first paragraph on page 16 the authors note that the analysis method they used to make indirect comparisons may not be familiar to all readers. This is certainly true and therefore it would be helpful if the authors could include a brief description of it. Also, it is not entirely clear what the authors mean when they say they conducted “adjusted” analyses.

Minor Essential

1. “all” should probably be dropped from the last sentence of the “Purpose” section on page 1 since only select new agents are evaluated.

2. The last sentence of the first paragraph on page 4 is incomplete

3. In the first paragraph of the “Results” section (page 9) the authors indicate that 8 abstracts were reviewed that were preliminary reports on five of the included studies; however nine references are given.

4. In the first paragraph on page 10 the authors indicate that five studies used interferon as the comparator; however in the first sentence of the “Meta-Analysis” section the number is given as six. In addition there is a statement that these studies were conducted predominantly in previously untreated patients; however from Table 1 prior treatment was an exclusion criterion.

5. The second paragraph of the “Discussion” indicates that seven papers were included in the study. The authors should clarify that five were manuscript publications and two were abstracts.

6. For completeness it would be helpful to include in Figure 2 the results of the individual study comparisons to placebo.
Discretionary

1. The interpretation of Figure 3 is not obvious and the figure does not add substantively to the manuscript. The authors should perhaps consider dropping it or provide an explanation of how to interpret it.

2. Table 3 is not referenced in the text. It is a very nice summary of adverse events; however it does not seem to add to the main focus of the paper. The authors should perhaps consider dropping it, or provide a brief discussion of the data.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.