National Institute for Health and Clinical Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Pazopanib for the first line treatment of advanced and/or metastatic renal cell carcinoma

Please find enclosed the ERG report prepared for this appraisal.

You are asked to check the ERG report from Aberdeen HTA Group to ensure there are no factual inaccuracies contained within it. If you do identify any factual inaccuracies you must inform NICE by **5pm, Friday 24 September 2010** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The attached proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

29th October 2009

Issue 1 OS HR sunitinib vs. IFN

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Executive Summary. Page IV For the comparison of sunitinib with IFN there was no evidence of a statistically significant difference between the groups (HR 0.821, 95% CI 0.673 to 1.001). These data were not used in the economic model, which used an exploratory analysis in the subset of patients who did not receive any post study cancer treatment (**HR 0.467** [95% CI 0.483 to 0.870]).	HR 0.647	The correct hazard ratio is 0.647 (Motzer 2009). It should be noted that, of the reported hazard ratios (HR) for OS for sunitinib versus IFN (ITT at interim analysis: HR 0.65 [0.45-0.94]; ITT at final analysis: HR 0.821 [0.673-1.001]; Censoring on cross-over: HR 0.808 [0.661-0.987), the one chosen for use in the indirect comparison and economic model was the most favourable towards sunitinib.

Issue 2 Objective Response Rate (IRC vs. INV)

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Executive Summary. Page IV In study VEG105192 overall response rate was higher for pazopanib (32%, 49/155) compared with placebo (4%, 3/78), p<0.001. For the comparison of sunitinib with IFN the response rate for sunitinib (47%, 176/375) was higher than that for IFN (12%, 46/375), p<001).	In study VEG105192 overall response rate was higher for pazopanib (32%, 49/155) compared with placebo (4%, 3/78), p<0.001. For the comparison of sunitinib with IFN the response rate for sunitinib (31%, 31/103) was higher than that for IFN (6%, 6/20), p<0.001) as reported in Motzer 2007 publication. Values reported above are based on assessment by Independent Review Committee (IRC)	This statement is misleading in not comparing like with like. The objective response rate (ORR) of 32% quoted for pazopanib in the VEG105192 trial is based on assessment by an Independent Review Committee (IRC). The updated 47% ORR quoted for sunitinib includes 88 scans based on investigators' assessment (Motzer 2009). The ORR for sunitinib by IRC assessment was 31% as reported in the Motzer 2007 publication.

Issue 3 Objective Response Rate (IRC vs. INV)

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Executive Summary. Page VII There was uncertainty surrounding the estimates reported by the indirect comparison, relating to the data used to derive the hazard ratios used to estimate relative effectiveness. While the pazopanib and sunitinib studies limited inclusion to participants with ECOG performance status 0 or 1, three of the IFN studies contained some participants with ECOG performance status 2 (i.e. a worse prognosis). This might make the relative performance of pazopanib and sunitinib against IFN appear better than it actually is.	There was uncertainty surrounding the estimates reported by the indirect comparison, relating to the data used to derive the hazard ratios used to estimate relative effectiveness. While the pazopanib and sunitinib studies limited inclusion to participants with ECOG performance status 0 or 1, three of the IFN studies contained some participants with ECOG performance status 2 (i.e. a worse prognosis). ***Theoretically***, this might make the relative performance of pazopanib and sunitinib against IFN appear better than it actually is. ***However, there is no evidence that the effects of treatment with pazopanib, sunitinib or IFN measured in terms of hazard ratios differ in subgroups of patients defined on the basis of performance status.***	While PFS and OS are likely worse for patients with worse performance status, there is no evidence that the *effects of treatment* with pazopanib, sunitinib or IFN measured in terms of hazard ratios differ in subgroups of patients defined on the basis of performance status—i.e., there is no evidence of effect modification. Worse performance status alone in the IFN trials is not sufficient in and of itself to bias the findings in favour of pazopanib and sunitinib.

Issue 4 Sensitivity analyses

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Executive Summary. Page VII The manufacturer concentrated on presenting a series of one-way sensitivity analyses which demonstrated that the cost-effectiveness results are not greatly altered by univariate changes. They did not consider the joint impact of changes in several parameters simultaneously. Furthermore, given the imprecise and potentially biased estimates of survival the deterministic analyses fail to fully illustrate the degree of uncertainty that exists.	The manufacturer concentrated on presenting a series of one-way sensitivity analyses which demonstrated that the cost-effectiveness results are not greatly altered by univariate changes. Furthermore, given the imprecise and potentially biased estimates of survival the deterministic analyses fail to fully illustrate the degree of uncertainty that exists.	The impact of changes in multiple parameters simultaneously is represented in the dispersion of results in the probabilistic sensitivity analyses.

Issue 5 Exclusion of CRECY study

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Section 4.1.2 p.10 The manufacturer's submission also excludes one interferon-α study (CRECY trial) by Negrier and colleagues[22] that was included in the systematic review (with the same inclusion and exclusion criteria as above) on the basis that "a non-immunotherapy control arm was not used in this study".	The manufacturer's submission also excludes one interferon-α study (CRECY trial) by Negrier and colleagues[22] that was included in the systematic review (with the same inclusion and exclusion criteria as above) on the basis that "a non-immunotherapy control arm was not used in this study". ***However, it is recognized that this study does not contain a non-active control arm which prevents its inclusion as part of the indirect comparison conducted by the manufacturer.***	It should be noted that the CRECY study could not be included in the adjusted indirect comparison since, as previously stated, it did not contain a non-active control arm to provide a 'bridge' to the VEG105192 study of pazopanib versus placebo/BSC as part of the indirect comparison.

Issue 6 Selection of baseline variables

Description of problem	Description of proposed amendment	Justification for amendment
ERG report – Section 4.1.6 p.20 These baseline variables were used across the analyses (including the alternative analyses described later in this section) with the rationale for their use merely stated as being based on prior literature and goodness-of-fit statistics, without referencing sources or presenting results of relevant analysis to justify the choice of these covariates.	These baseline variables were used across the analyses (including the alternative analyses described later in this section) with the rationale for their use merely stated as being based on prior literature and goodness-of-fit statistics.	This is not quite correct. The baseline variables used to adjust our analyses were selected on the basis of clinical evidence, clinical opinion and availability of data in the VEG105192 study. Furthermore, all the variables included are well know predictors of mortality in the general population (age and sex) and/or in patients with advanced RCC (Motzer score, time since diagnosis, stage at diagnosis, presence of liver metastases, and number of metastatic sites). One of the variables selected was the Memorial Sloan Kettering Cancer Center (MSKCC) risk score. This is a widely accepted and validated predictive tool for survival in advanced RCC (Motzer 1999). Time since diagnosis, stage of disease at diagnosis, and number of metastatic sites are all well documented prognostic factors in advanced RCC. The reference sources cited in our

		submission were review articles summarising the available evidence (Bukowski 2009; Furniss 2008; Elson 1988). In response to a clarification question from the ERG regarding the rationale for the inclusion of 'presence of liver metastases', GSK cited the Negrier 2005 study as additional evidence as well as the fact that pazopanib-treated patients with liver metastases in the VEG105192 trial had a median PFS of 5.6 months compared with 12.9 months for patients without liver metastases (p=0.005). *Please see references at the end of the document*

Issue 7 Statistical analyses undertaken – of Test statistics

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 4.1.7 p.22 The manufacturer defines the upper limit of the 95% confidence interval of the causal rate ratio, described as being based on the inversion of the test statistic, as the largest parameter value for which the p-value is greater than 0.05. The inversion of the test statistic presents a more conservative value for the causal rate ratio but was not used in the base case analysis, despite these upper limits being reported. In the weighted RPSFT analysis, this value is stated as being -0.05, which represents a very small beneficial effect for pazopanib. If this had been used in the analysis, the result of the RPSFT model would be very similar to the results of the ITT analyses.	Academic experts consulted seem to agree that the ERG statement does not coincide with the theoretical aspects behind the two methods employed and the results derived from them.	In the document we propose two alternative methods for calculating the confidence interval on psi (and by implication the causal rate ratio [CRR]=exp(psi)): (1) based on inversion of the test statistics-i.e., the largest parameter values for which the p-value is greater than 0.05; and (2) Bootstrapping. The upper limit of the confidence interval on psi obtained by inversion of the test statistic is lower (less conservative) than that obtained by bootstrapping. It is necessarily true that the upper limit of the CI (using either method) is greater than the estimate of central tendency, and in that sense, if the HR corresponding to the upper limit of the CI were used in the model it would be conservative. However, the inversion of the test statistic does not provide an estimate of psi *per se*.

Issue 8 Statistical analyses undertaken – Inversion of Test statistics

Description of problem	Description of proposed amendment	Justification for amendment
Another reason to support the presentation of results based on the inversion of the test statistic is that it could be a solution when there is a multimodal p-value distribution. In the example of the unweighted unadjusted analysis, the p-value distribution plot indicates that the causal rate ratio would be approximately +1.3, which therefore would not indicate any significant evidence that cross-over was masking some beneficial effect of pazopanib.		As note in the Justification for Issue 6, the psi value obtained based on inversion of the test statistic represents the upper confidence limit of the estimate of psi, not an estimate of psi *per se*.

Issue 9 Statistical analyses undertaken - RPSFT

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 4.1.7 p.23 However, the RPSFT method is heavily weighted towards the early follow-up period and the analysis only controlled for cross-over from placebo to pazopanib and not receipt of other post-study anti-cancer therapies, particularly important as there was an imbalance between the groups with more pazopanib patients (24%) receiving anti-cancer therapies compared with placebo patients (12%).	However, the RPSFT method is heavily weighted towards the early follow-up period and the analysis only controlled for cross-over from placebo to pazopanib and not receipt of other post-study anti-cancer therapies, particularly important "as there was an imbalance between the groups with more pazopanib patients (24%) receiving **other** anti-cancer therapies (**excluding pazopanib)** compared with placebo patients (12%).	12% is quoted as the percentage of subjects who took non-pazopanib systemic treatments. However this is actually the percentage of subjects who took only non-pazopanib systemic treatments and excludes the subjects who took both pazopanib and other non-pazopanib systemic treatments. To compare the true effects of not adjusting for the effects of non-pazopanib systemic treatments one needs to account for all subjects who received a non-pazopanib systemic treatment. In the placebo group, 8 subjects or 10% took pazopanib and an additional treatment (sunitinib, interferon, temsirolimus or everolimus). This brings the total who received additional non-

		pazopanib systemic treatments to 22%. When this 22% is compared to the 24% on the pazopanib arm, it is hard to argue that this 2% difference would result in a substantial bias in favor of the pazopanib arm. Percentages can be derived from table 1.3 (OS final addendum – 20 July 2010) Further if one examines the actual treatments received by both arms, it seems that the placebo arm might have overall been receiving a more active mix of treatments. The placebo arm seems to have gotten more access to *m-tor* treatments. There is a known survival benefit for treatment with temsirolimus in poor risk treatment naïve subjects. Analyses to adjust for crossover have also demonstrated a likely survival benefit for everolimus in TKI failures. In contrast the pazopanib arm seems to have received more cytokine treatment and more treatment in the "Other" category. The Other category includes treatments such as Vinblastine, MPA and thalidomide, which we have already argued are not effective treatments. None of these 4 subjects experienced a partial response in relation to these Other treatments. There were similar percentages with access to sunitinib, sorafenib, and bevacizumab. Finally, while it is correct that the RPSFT analysis did not control for receipt of other anti-cancer therapies this likely did not materially bias the findings in favour of pazopanib as results from IPCW analyses were more favourable when receipt of other anti-cancer therapies were considered informative censoring events vs. when they were not.

Issue 10 Statistical analyses undertaken - RPSFT

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 4.1.7 p.24 Overall, the manufacturer has presented a set of analyses which comprehensively covers the range of methodologies available to adjust for cross-over. However, care should be taken when assessing trials that have used relatively new methods as there is no consensus on the best approach to use and these methods still require further development. In this particular analysis, the results used for the base case economic model utilise a new methodology which is still to be peer-reviewed and published. Also, the rationale for deriving the causal rate ratio is unclear and there potentially could be a large underestimation of the hazard ratio as a result	Overall, the manufacturer has presented a set of analyses which comprehensively covers the range of methodologies available to adjust for cross-over. However, care should be taken when assessing trials that have used relatively new methods as there is no consensus on the best approach to use and these methods still require further development.	The theory of RPSFT has been peer reviewed and used a number of times in applications. The use of the optimal weighted estimator has been peer reviewed (appendix 4 p 279-281 of Robins 1993 in book Aids methodology) although it has not before been used in an application The estimation of the causal rate ratio given an estimate of psi can be shown to be valid via an elementary argument so the claim of potential underestimation given an estimate of psi is incorrect. NB. Please see attached file which contains comments from Professor Jamie Robins on this particular point.

Issue 11 Statistical analyses undertaken - RPSFT

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 5.2.2 p.68 As noted in Chapter 4, several methods were used to estimate hazard ratios and a key concern is the derivation of the causal rate ratio for the weighted unadjusted analysis used in the base case analysis. These concerns are expanded upon in Section 4.1.7 but briefly the	As noted in Chapter 4, several methods were used to estimate hazard ratios and a key concern is the derivation of the causal rate ratio for the weighted unadjusted analysis used in the base case analysis. These concerns are expanded upon in Section 4.1.7	As psi gets very large or small artificial censoring goes to one so power is essentially zero because of the few cases that are not artificially censored in either one or the other treatment arm and thus p values cannot be significant. Fortunately the large positive and negative values of psi with very poor power are usually biologically implausible (as they

statistical model does not appear to have a single likeliest value and where a choice has been made it has been made in favour of pazopanib. .		would imply the drug increases or decreases survival times by a large factor. Thus we did not include extreme values of psi as they provide no relevant information. It should also be noted that the grid plots for the unweighted and weighted unadjusted analyses are fundamentally different. With the unweighted analysis, there is a large mass under the curve with multiple peaks ranging from -2 to 0 whereas for the weighted analysis there were two distinct peaks, with one at the biologically implausible value of 2.7 (suggesting a CRR of 15) Please see attached document from Dr Robins.

Issue 12 Acquisition cost for sunitinib

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 5.4.2 p.96 When deciding on the cost of pazopanib the manufacturer stated that the list price of pazopanib has been set at parity with the sunitinib list price (calculated on a price per day basis). Assuming that the HR data for OS suggest that there may be no differences between pazopanib and IFN, that is HR>1 for overall survival (strictly speaking there is no evidence of a difference and the confidence interval for this estimate is very wide) the total cost of pazopanib is £34,647 and the drug acquisition cost is £27,476. On the other hand the total cost of sunitinib is £36,179 and the drug acquisition cost **is £28,956**. The rebate that is being offered is slightly more than that	When deciding on the cost of pazopanib the manufacturer stated that the list price of pazopanib has been set at parity with the sunitinib list price (calculated on a price per day basis). Assuming that the HR data for OS suggest that there may be no differences between pazopanib and IFN, that is HR>1 for overall survival (strictly speaking there is no evidence of a difference and the confidence interval for this estimate is very wide) the total cost of pazopanib is £34,647 and the drug acquisition cost is £27,476. On the other hand the total cost of sunitinib is £36,179 and the drug acquisition cost is £28,**8**56. The rebate that is being offered is slightly more than that required to get the drug cost equal to sunitinib.	Drug acquisition cost for sunitinib should read £28,856

required to get the drug cost equal to sunitinib.		

Issue 13 Routine follow-up Cost

Description of problem	Description of proposed amendment	Justification for amendment
ERG report - Section 5.6 p.100 The other issue relates to routine follow-up costs used in the model. In the model it was assumed that as soon as someone progressed they stopped treatment. However, the cost estimates used in the model are based on the assumption that they are incurred e.g. follow-up costs are cited as £146 per month.	Routine follow-up costs used in the model are £146 per month of PFS and £228 per months of PPS.	Comment does not accurately reflect the cost inputs in the model

References

Bukowski RM. Prognostic factors for survival in metastatic renal cell carcinoma. Cancer 2009; May 15: 2273-2281.

Elson PJ, Witte RS, Trump DL. Prognostic factors for survival in patients with recurrent or metastatic renal cell carcinoma. Cancer Res 1988; 48: 7310-7313.

Furniss D, Harnden P, Ali N, *et al.* Prognostic factors for renal cell carcinoma. Cancer Treat Rev 2008; 34: 407-426.

Motzer RJ, Mazumdar M, Bacik J, *et al*. Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma. J Clin Oncol 1999; 17: 2530-2540.

Motzer RJ, Hutson TE, Tomczak P, *et al.* Sunitinib versus interferon alfa in metastatic renal cell carcinoma. New Engl J Med 2007; 356: 115-24.

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.

Pyrhonen S, Salminen E, Ruutu M, Lehtonen T, Nurmi M, *et al.* Prospective randomized trial of interferon alfa-2a plus vinblastine versus vinblastine alone in patients with advanced renal cell cancer. J Clin Oncol 1999; 17(9): 2859-2867.

Negrier S, Gomez F, Douillard JY, et al. Groupe Francais d'Immunotherapie. Prognostic factors of response or failure of treatment in patients with metastatic renal carcinomas treated by cytokines: a report from the Groupe Francais d'Immunotherapie. World J urol 2005; 23 (3): 161-5.