EVEROLIMUS FOR THE SECOND-LINE TREATMENT OF ADVANCED AND/OR METASTATIC RENAL CELL CARCINOMA

APPEAL AGAINST THE FINAL APPRAISAL DETERMINATION ISSUED BY THE NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE ON 19 NOVEMBER 2010

EXECUTIVE SUMMARY

1. Ground 1: Procedural Unfairness

   1.1 NICE's failure to disclose to Novartis the modified economic model upon which its guidance is based, lacks transparency and is unfair. As a consequence of this:

   - Novartis has been unable to investigate the apparent inconsistency in the estimates of overall survival (OS) used at various parts of the assessment.
   - Insufficient information has been provided to allow Novartis to investigate or verify a criticism made by the ERG regarding the calculation of the probability that the ICER is below £50k/QALY using Novartis' probabilistic sensitivity analysis.
   - Novartis was unable to evaluate appropriately the uncertainty surrounding the ERG's extrapolation of OS using Weibull curves and calculation of the transition probabilities without access to the modified model or adequate information from the ERG.

   1.2 The lack of transparency in relation to the extrapolation of OS data associated with everolimus therapy using a Weibull curve is unfair.

   1.3 The lack of opportunity afforded to Consultees to scrutinise and comment on the ERG's "exploratory" analyses (modifications to Novartis' model) which form the basis of the recommendations in the FADs dated June 2010 and November 2010 constitutes procedural unfairness.

2. Ground 2: Perversity

   2.1 The reasons given by the Appraisal Committee for refusing to consider the investigation of uncertainty surrounding the hazard ratio for overall survival (OS) based on a more clinically plausible range, carried out by Novartis, are inconsistent with the request from NICE to investigate a plausible range, and the evidence. This therefore constitutes perversity.

   2.2 The approach of the Appraisal Committee to the possibility of uncertainty in the assessment of cost effectiveness is inconsistent with that followed in other appraisals and is therefore perverse.

2.3 Due to the heterogenous nature of the studies and the patient populations included in the Delea meta-analysis referred to in sections 3.6, 4.5 and 4.10 of the FAD, dated November 2010, reliance of the Appraisal Committee on the ratio of PFS:OS of 1:1.4 to quantitatively justify the survival gain of 5.9 months from the ERG's analysis is perverse. The results from this analysis are unlikely to represent so specifically the OS gain conferred by everolimus.

2.4 The Appraisal Committee has disregarded the available evidence for OS in patients who receive BSC

2.5 The estimates of mean OS associated with BSC alone, relied upon by the Appraisal Committee are inconsistent and do not reflect the referenced calculations of the ERG.

2.6 Reliance of the Appraisal Committee on a mean probabilistic ICER to justify the decision not to recommend everolimus is perverse as, due to the nature of the analysis, the mean probabilistic ICER will vary from one run to another.

INTRODUCTION

Everolimus (Afinitor) is indicated for the treatment of patients with advanced renal cell carcinoma, whose disease has progressed on or after treatment with VEGF-targeted therapy. It was designated an orphan medicinal product by the European Commission in the indication of renal cell carcinoma, on 5 June 2007. Everolimus is the subject of a marketing authorisation granted to Novartis by the European Commission under the centralised procedure on 3 August 2009, following a favourable opinion by the CHMP on 29 May 2009. It is supplied in the UK by Novartis Pharmaceuticals (UK) Limited as Afinitor

Everolimus is a derivative of rapamycin and acts specifically to inhibit mTOR a protein involved in the regulation of tumour cell division and blood vessel growth in numerous human cancers.

Renal cell carcinoma (RCC) is the most common primary renal malignant neoplasm in adults and the eighth most common cancer in England and Wales. It accounts for approximately 90 per cent of renal tumour malignancies and 3 per cent of all new cases of cancer diagnosed in men and just over 2 per cent of cancers in women in the UK (excluding non-melanoma skin cancer).

RCC is often asymptomatic until it reaches a late stage and 25 - 30 per cent of patients have metastatic disease at the time of presentation. Approximately 44 per cent of patients diagnosed with RCC in England and Wales live for at least 5 years after initial diagnosis and about 40 per cent live for at least 10 years. However, metastatic disease is associated with a poor prognosis and approximately 90 per cent of patients diagnosed with metastatic RCC have died by five years after diagnosis. Without treatment, patients with advanced RCC have a median survival of less than 12 months.

Until recently, the standard treatment for RCC in England and Wales was radical nephrectomy and cytokine therapy with interleukin-2 or interferon-alpha. In March 2009, NICE approved sunitinib for first line use in patients with advanced RCC who are suitable for immunotherapy and have good performance status (TAG 169). However, there is currently

no NICE recommended treatment for patients with advanced RCC who do not respond to 1st line VEGF-targeted therapy with sunitinib, or whose disease progresses whilst on treatment.

PROCEDURAL HISTORY OF THIS APPRAISAL

The single technology appraisal of everolimus commenced in 2009 and the final scope was issued in July that year.

On 30 September 2009, Novartis provided its submission in relation to everolimus for the treatment of advanced RCC, to NICE. This submission included an electronic version of the economic model which formed the basis for Novartis' assessment of cost effectiveness.

The Peninsular Technology Assessment Group (PenTag) was instructed to act as the Evidence Review Group (ERG) in this appraisal. On 16, 22 and 26 October 2009, following preliminary work by PenTAG, the Institute wrote to Novartis requesting further clarification relating to the clinical and cost effectiveness data. Novartis responded to the requests for clarification in a further submission on 30 October 2009. PenTAG completed their ERG report on 30 November 2009 and, on 3 December, a copy of the report was provided to Novartis, to confirm its accuracy.

At this stage, Novartis became aware of the ERG's preference for a Rank Preserving Structural Failure Time (RPSFT) statistical analysis to address the issue of cross over in the principal RECORD-1 clinical trial investigating the safety and efficacy of everolimus. On 10 December 2009, Novartis therefore provided comments on the ERG report together with a formal request to submit a further analysis using the RPSFT approach to trial data, as opposed to the Inverse Probability Censoring Weights (IPCW) approach contained in its original submission. NICE agreed and the additional analysis was submitted to the Institute on 18 December 2009, together with confirmation from the Department of Health that a patient access scheme (PAS), proposed by Novartis, had been approved.

The first meeting of the Appraisal Committee took place on 13 January 2010. Following that meeting, an appraisal consultation document (ACD) was sent to Novartis on 2 February 2010. At section 1 of the ACD, the Appraisal Committee's preliminary recommendations were that:

> "Everolimus is not recommended for the second-line treatment of advanced renal cell carcinoma".

Novaratis' response to the ACD was sent to NICE on 2 March 2010, together with additional evidence and analyses prompted by the ERG report and the preliminary guidance.

The Appraisal Committee met for a second time on 9 March 2010 and considered the consultation responses. They agreed however that a third meeting of the Committee would be required to complete this review. On 12 March 2010, the ERG requested additional analyses and these were provided by Novartis on 24 March 2010.

The third meeting of the Appraisal Committee took place on 12 May 2010, following which a final appraisal determination (FAD) dated June 2010, was sent to consultees.

On 19 July 2010, Novartis provided an updated patient access scheme to NICE. On 11 August 2010, a fourth meeting of the Appraisal Committee took place to consider the revised

patient access scheme. Following this meeting, the Institute advised consultees and commentators that further analyses were needed by the Appraisal Committee to clarify uncertainty associated with new evidence on the total costs of use of everolimus in the NHS. Novartis provided the requested analyses on 24 September 2010.

On 13 October 2010, a fifth meeting of the Appraisal Committee took place. The second FAD was released to commentators and consultees on 19 November 2010. The proposed guidance at section 1.1 of the second FAD was unchanged from that contained in the ACD and previous FAD, dated June 2010.

GROUNDS OF APPEAL

1. Ground 1: Procedural Unfairness

1.1. NICE's failure to disclose to Novartis the modified economic model upon which its guidance is based, lacks transparency and is unfair

A copy of the adjusted model or clarification regarding the ERG's adjustments was requested from NICE on three occasions, on 30 June 2010, 17 August 2010 and 23 November 2010. However, a copy of the adjusted model has never been provided and the responses to our requests for clarification have not been adequately addressed.

In the context of this single technology appraisal, the economic model relied upon by the Appraisal Committee was prepared by Novartis and submitted in electronic format with Novartis' original submission in September 2009. The analysis included in this submission was based on the IPCW approach. Subsequently, on the 18 December 2009, an alternative analysis based on the RPSFT statistical approach was submitted in response to comments from the ERG that a comparison of the results from the IPCW and RPSFT approach would be of interest.

A fully executable copy of the model was supplied to the ERG, who made various modifications to both the IPCW and RPSFT analyses reported (a) as described at Section 6.2.1 of the ERG Report, dated 30 November 2009 (IPCW analysis) and (b) as described on page 2, of the ERG Report, dated 21 December 2009 (RPSFT analysis).

Further analyses were submitted by Novartis in March 2010, in response to the ACD, which had become available from the RECORD-1 trial. The updated RPSFT analysis was then modified by the ERG who produced a base case ICER of £58,316/QALY for everolimus therapy. Henceforth we will refer to ERG's modifications to the Novartis model as the ERG adjusted model. Following receipt of the first FAD in June 2010, Novartis considered an appeal, but felt that offering a revised PAS would facilitate faster patient access to treatment. We therefore offered to reduce the cost of everolimus further within a revised PAS.

On the 30 June 2010 we requested a copy of the ERG adjusted model to incorporate the revised PAS but were instead provided with the ERG's transition probabilities in order that we could replicate the ERG's ICER of £58,316/QALY. The revised PAS was incorporated into Novartis' replication of the ERG's modified model and submitted by Novartis on 19 July. At the request of the Appraisal Committee, a probabilistic sensitivity analysis was submitted by Novartis on 17 August 2010. The

ERG made further modifications to the probabilistic analysis as described in their report dated 7 October 2010.

A copy of the probabilistic model as adjusted by the ERG was requested by Novartis in an email dated 23 November 2010. Novartis had sought to understand the changes made by the ERG to its probabilistic economic model. However, in certain important respects (including calculation of the probabilities that everolimus would be cost-effective at values of £30,000 and £50,000) it was unable to do so based on the information and explanations provided. Novartis could not therefore investigate appropriately the conclusions reached by the ERG and relied upon by the Appraisal Committee (see e.g. section 4.12 of the FAD dated November 2010) in formulating its conclusions regarding everolimus. The lack of adequate explanation or access to the modified version of Novartis' economic model has precluded Novartis understanding the assessment of cost-effectiveness by the Appraisal Committee including in relation to the matters set out below:

- Novartis has been unable to investigate the apparent inconsistency in the estimates of overall survival (OS) used at various parts of the assessment.

  There are a number of inconsistencies between the estimates of OS accepted by the Appraisal Committee and reported in the FADs and those arising from Novartis' replicated model despite the ICERs from both models being identical ie £58,316/QALY. However, lack of disclosure of the ERG's version of the model has prevented Novartis investigating the source of the discrepency.

  Following release of the original FAD on the 25 June 2010, Novartis requested access to the ERG modified version of the model from NICE, in a telephone conversation with Meindert Boysen on 30 June 2010, in order to evaluate the impact of the revised PAS. However, instead of receiving the adjusted model, on 1 July, Novartis was provided with a brief description of the ERG's modifications to the Novartis model together with the ERG's new transition probabilities. These transition probabilities were said to be based on the ERG's use of a Weibull curve to extrapolate OS. This information allowed us to replicate, using our model, the base case ICER of £58,316/QALY calculated by the ERG (reported at section 3.24 of the original FAD. This replicated model was then used to evaluate the impact of the revised PAS giving an amended ICER of £49,300 referred to in Section 4.17 of the FAD dated November 2010. This ICER of £49,300/QALY was accepted by the Appraisal Committee as being the correct deterministic ICER for everolimus with the revised PAS (section 4.11 of the FAD, dated November 2010).

  However, although we had been able to replicate the ERG's base case ICER of £58,316/QALY, reported at paragraph 3.24 of the original FAD, using our model, the estimates of survival producing this ICER and subsequently the ICER of £49,300 with the revised PAS, in our replicated model, are not consistent with the estimates of survival accepted by the Appraisal Committee in the FAD, dated November 2010. In particular, our estimates of OS for everolimus plus BSC, OS for BSC and the difference between these estimates (ie the improvement in OS conferred by everolimus) which give rise to the ICER of £58,316/QALY (and subsequently £49,300 with the revised PAS) are

different to those reported to underpin the ERG's base case ICER. Section 4.7 of the FAD, dated November 2010, states that the Appraisal Committee accepted that the mean OS for everolimus plus BSC was 16.7 months, for BSC alone the mean OS was 10.8 months, and the difference in OS was 5.9 months. When we replicated the ERG's ICER of £58,316/QALY using our economic model, we found that this figure was underpinned by a mean OS for everolimus plus BSC of 14 months, a mean OS for BSC alone of 8.9 months, and a difference in OS of 5.2 months.

In summary, the ERG's calculation of a base case ICER of £58,316/QALY (subsequently £49,300/QALY with the revised PAS) using the estimates for OS accepted by the Appraisal Committee, are wholly unclear and may not be investigated by Novartis without access to the modified version of the economic model.

- Insufficient information has been provided to allow Novartis to investigate or verify a criticism made by the ERG regarding the calculation of the probability that the ICER is below £50k/QALY using Novartis' probabilistic sensitivity analysis.

  At paragraph 3.31 of the FAD dated November 2010, the Appraisal Committee notes that the ERG stated that it had identified an error in the probabilistic sensitivity analysis carried out by Novartis and described at paragraph 3.28 of the FAD. The ERG had said "that simulations that resulted in dominated outputs were included when the ICER threshold was set at zero, suggesting that it is possible for everolimus plus best supportive care to be cheaper than best supportive care alone". The ERG stated that it re-ran the analysis "and corrected for the error identified" (paragraph 3.32 of the FAD dated November 2010) and those results were relied upon by the Appraisal Committee in deciding not to recommend everolimus. Novartis did not believe the ERG's criticism was correct and attempted to investigate the conclusions set out in the FAD. In the absence of the ERG's adjusted model Novartis attempted to explore this further using our own model. However despite repeatedly re-running the model with 1,000 iterations, we were unable to identify any outputs where everolimus plus BSC was less costly than BSC alone as implied by the FAD. Therefore, in the absence of the ERG's modified version of the model, Novartis has been unable to verify or adequately investigate the conclusions of the ERG.

- Novartis was unable to evaluate appropriately the uncertainty surrounding the ERG's extrapolation of OS using Weibull curves and calculation of the transition probabilities without access to the modified model or adequate information from the ERG.

  Only limited information was provided by the ERG in relation to its extrapolation of OS using Weibull curves and calculation of the associated transitions probabilities in its report dated 27 April 2010. Some additional limited information was provided on 1 July in response to our request for the ERG's adjusted. Novartis was concerned about the way the Weibull curves and associated transition probabilities were derived as we were unable to reproduce the Weibull curves from the information provided by the ERG.

> Accordingly, on the 17 August 2010, Novartis requested clarification of some technical points regarding the modifications to the model made by the ERG, in order to conduct the probabilistic sensitivity analysis requested by the Appraisal Committee. This request included clarification regarding how the transition probabilities were derived from the Weibull curves in the ERG adjusted model and information on the 95% confidence intervals (CI) for the Weibull curves. The ERG's response was provided to Novartis on 1 September 2010. However the information provided did not adequately address the questions we had asked and the ERG said that they could not locate their original workings in order to provide the CI around the Weibull curves.
>
> The failure by NICE or the ERG to disclose either the modified version of the model or sufficient information to enable us appropriately to understand the ERG's use of the Weibull curves in this appraisal has prejudiced us in our ability to investigate the extrapolation of OS by the ERG, including (as noted by the ERG in its report of 4 August 2010) the uncertainty around the Weibull curves. Section 3.31 of the FAD, dated November 2010 states, "The ERG was unsure whether all sources of uncertainty had been included in the probabilistic sensitivity analysis (for example, no evidence could be found by the ERG that uncertainty around the Weibull survival curve had been incorporated.)". However, failure of the ERG to provide the 95% CI around the Weibull curves required Novartis to seek to explore uncertainty in relation to OS using a different approach namely by varying the transition probabilities. Therefore the criticism that we have failed to explore uncertainty around the Weibull survival curve is unfair as we were not provided with the requested information to undertake this approach.

NICE's Guide to the Single Technology Appraisal Process (the STA Guide) states at paragraph 3.5.29 that "*if the manufacturer or sponsor has submitted an economic model, NICE offers to send it (in its executable form) to consultees and commentators during consultation on the ACD (if produced) or with the FAD*". It is implicit that the version of the model that should be disclosed in accordance with NICE's procedures is that relied upon by the Appraisal Committee for the purposes of its recommendations set out in the FAD. This was confirmed by the Court in R ota Bristol-Myers Squibb Pharmaceuticals Ltd -v- NICE (2009)

> "I*f the ERG change the manufacturer's model by inputting fresh data into it and the statistical calculations derived from the revised model form the basis of its written conclusions, then fairness would tend to require that consultees who sought to challenge the revised cost effectiveness figures should have access to the information that lead to the conclusions in the report. If that information can only properly be supplied by handing to the manufacturer the fully executable model as modified by the ERG, then the position would be materially the same as that pertaining in the Eisai case*".

Accordingly, in all the circumstances of this case, the refusal by NICE to disclose the modified version of Novartis' economic model, has prejudiced Novartis in its ability to understand, investigate and comment upon the conclusions reached by the Appraisal Committee, and is, accordingly, procedurally unfair.

1.2. The lack of transparency in relation to the extrapolation of data on OS associated with everolimus therapy using a Weibull curve is unfair

The key factor in determining the cost effectiveness of everolimus is the estimate of the additional survival benefit associated with treatment (section 4.8 of the FAD, dated November 2010). In circumstances where OS may not be determined from the RECORD-1 clinical trial as a result of the high proportion of patients who crossed over to the everolimus arm, it is necessary to estimate OS through modelling. In this context, the Appraisal Committee accepted the ERG's method of extrapolation of data using a Weibull curve, rather than the alternative method proposed by Novartis (section 4.9 of the FAD dated November 2010).

Novartis had some concerns regarding the use of the Weibull curve by the ERG, to calculate OS for the purposes of the economic model. In particular, the mean estimate of OS for BSC patients obtained from this approach reported in section 4.7 of the FAD, dated November 2010, as 10.8 months exceeds:

a) that observed in clinical practice, based on an audit of patients progressing on sunitinib treatment and provided to NICE with our ACD response of 2 March 2010 - 2 to 5 months;

b) the views of treating clinicians solicited from a survey, also provided to NICE with our ACD response of 2 March 2010 - 6 months; and

c) the estimate of OS put forward by the Royal College of Physicians in its response to the ACD (4 to 5 months).

Furthermore, it was not possible for Novartis to investigate adequately the level of uncertainty surrounding the extrapolation of data by the ERG using the Weibull curves, as the 95% confidence intervals around the Weibull curves were not provided and insufficient information was supplied regarding the derivation of the transition probabilities from the Weibull curves.

As a result of its inability to understand adequately and to investigate the extrapolation of OS data carried out by the ERG, on 17 August 2010, Novartis requested further information regarding the ERG's use of Weibull curves in this appraisal, including the 95% confidence intervals associated with the ERG's calculations. By email dated 1 September 2010, Novartis received the ERG's response, answering some of Novartis' queries but declining to provide the 95% confidence intervals, or details of the analysis, on the basis that they "could not locate [their] original workings". The ERG stated that "the curves represent the best fit to the data using regression analysis which can easily be repeated to obtain these values".

Novartis were unable to reproduce the ERG generated Weibull curves using the information provided by the ERG and the full dataset from the RECORD-1 trial. On further investigation it was found that the Weibull curves could seemingly be replicated if the last data point (from cycle 6 of the RECORD-1 trial) was omitted. Although no reference is made to this omission in the FAD, dated November 2010 it is briefly discussed in the ERG Report dated 27 April 2010. This raises two issues:

- Firstly section 4.9 of the FAD, dated November 2010 states that, "The Committee accepted that the use of a Weibull distribution was a more appropriate method for fitting and extrapolating the curve, as all available data was used."

  However, if Novartis could only replicate the ERG's Weibull curves if the last data point were excluded, then it does not appear to be the case that all available data were used and the Appraisal Committee acted on incorrect information when it concluded that the Weibull curve fitting by the ERG was a more appropriate method for extrapolating data in this appraisal.

- Secondly, if, despite the Appraisal Committee's statement at section 4.9 of the FAD, dated November 2010, the final data point has been omitted from the Weibull curves produced by the ERG, then the failure by the ERG to explore the impact of this omission is procedurally unfair. From Novartis own exploratory analysis we know that the inclusion of these data within the analysis will decrease the estimate of mean OS for BSC alone patients and thus reduce the deterministic ICER by around £3,000 bringing the ICER down to around £46k/QALY. While we have been unable to understand appropriately the use of Weibull curves by the ERG, if they have omitted the final data point from the RECORD-1 trial if the AC has relied upon an extrp of OS which omits the data point from REC this must be justified and failure to do sop is perverse. In addition, this is likely to reduce the mean probabilistic ICER and improve the probability that the mean probabilistic ICER will be below £50k/QALY. However without access to full explanations of the approach followed by the ERG, including the 95% CI for the Weibull curves and a description of how the transition probabilities were calculated, Novartis is unable to investigate this issue fully and appropriately.

In summary, Novartis is entitled to know the detailed methods, data employed and resulting figures calculated by the ERG which were used for the purposes of its own evaluation of everolimus so that the ERG's conclusions may be properly investigated and tested. Clearly the full workings of the ERG should have been published with its reports or, at the least, retained and, if the ERG is now unable to produce or replicate its figures, this calls into question the validity of its conclusions. The situation whereby a key driver of the assessment of cost effectiveness, in an area where there was disagreement between the manufacturer and the ERG, cannot be investigated or tested, plainly prejudices the ability of consultees to understand and test the conclusions expressed in the FAD and is therefore unfair.

While we have been unable to understand appropriately the use of Weibull curves by the ERG, if the Appraisal Committee has relied upon an extrapolation of OS which omits the data point from the RECORD-1 trial, this should be justified and failure to do so is perverse.

1.3. The lack of opportunity afforded to consultees to scrutinise and comment on the ERG's "exploratory" analyses (modifications to Novartis' model) which form the basis of the recommendations in the FADs dated June 2010 and November 2010 constitutes procedural unfairness.

Following the ACD, Novartis submitted new analyses based on longer term data from the RECORD-1 trial to assist the Appraisal Committee. The ERG reviewed Novartis' submission and prepared a supplementary report dated 27 April 2010. This report introduced, for the first time, the use of Weibull curves to extrapolate OS, rather than the methodology which had been used by Novartis. In circumstances where the estimates of OS are a key driver of cost-effectiveness in this appraisal, this new approach by the ERG was of fundamental importance.

Novartis had no opportunity to review or comment upon the ERG's report of 27 April 2010 before this was considered by the Appraisal Committee, who then proceeded to issue the original FAD, which relied upon the ERG's conclusions.

In view of the importance of the new work carried out and presented by the ERG in its report of 27 April 2010, we believe it was unfair for the Appraisal Committee to issue a FAD before Novartis and other consultees had been permitted to consult in relation to this material. Instead, NICE should either have requested consultation on the ERG report before the meeting of the Appraisal Committee on 12 May 2010 or should have issued a second ACD for consultation following that meeting, as envisaged by paragraph 3.5.35 of NICE's Guide to the Single Technology Appraisal Process, rather than a FAD.

While the Appraisal Committee met again on 11 August and 13 October 2010, there was no opportunity for consultation on the supplementary reports prepared by the ERG and Novartis was permitted only to provide sensitivity analyses investigating uncertainty, rather than to make submissions in relation to the overall approach of the ERG.

For completeness, this appeal does not correct the lack of fairness inherent in the absence of consultation described above, as the Appeal Panel does not have power to reconsider the merits of any conclusions previously reached by the Appraisal Committee, unless these reach the perversity threshold.

2. Ground 2: Perversity

2.1 The reasons given by the Appraisal Committee for refusing to consider the investigation of uncertainty surrounding the hazard ratio for overall survival (OS) based on a more clinically plausible range, carried out by Novartis, are inconsistent with the request from NICE to investigate a plausible range and the evidence. This therefore constitutes perversity.

At paragraph 4.11 of the FAD, the Appraisal Committee refers to the probabilistic sensitivity analysis carried out by Novartis using adjusted 95% confidence intervals around the hazard ratio for OS associated with everolimus therapy. The Appraisal Committee however declined to consider this analysis stating that data relied upon by Novartis "were likely to be biased". The Committee's reasons for this conclusion are, however, inconsistent with the evidence and therefore perverse.

As explained above, the key factor determining cost effectiveness in the context of this appraisal is the calculation of OS associated with everolimus therapy plus best supportive care (BSC), compared with BSC alone. As result of patient crossovers to

the everolimus arm in the principal RECORD-1 clinical trial, the trial data on OS are confounded and it was necessary to estimate OS in the BSC arm by means of modelling. This included a recalculation of the hazard ratio (the probability of death or progressive disease at any stage) as well as the extrapolation of data referred to at point 1.2 above. The approach to recalculation of the hazard ratio preferred by the ERG was the Rank Preserving Structural Failure Time (RPSFT) method and this was the approach accepted by the Appraisal Committee. However, the 95% confidence intervals produced following this calculation were wide. Results from the RPSFT method are expressed as relative survival with a 1.93 fold longer survival in the everolimus plus BSC arm than the placebo plus BSC arm (95% CI: 0.5 – 8.5) giving a hazard ratio of 0.489 with a 95% confidence interval of 0.06 to 1.63.

While the 95% confidence intervals produced by the RPSFT methodology are statistically valid the extreme values are clinically implausible. By way of example, a relative survival of 0.5 would mean that patients treated with BSC alone would live twice as long as patients on everolimus treatment. There is no evidence to suggest that such an assumption may be valid and it is inconsistent with the available trial data. However by including these extreme possibilities in the analysis, the true uncertainty may be obscured. In a communication from NICE dated 16 August 2010 it was acknowledged that a probabilistic sensitivity analysis might be more complex due to the nature of the RPSFT method and therefore Novartis was asked to consider plausible ranges: "We appreciate there may be additional complexities in conducting such sensitivity analyses due to the constraints of the RPSFT method used to derive the estimates of OS and as such plausible ranges should be investigated."

Novartis therefore defined a range which was at least clinically plausible (albeit, we believe conservative), based on the available evidence for OS in patients treated only with BSC. For the upper limit of the range, data from the RECORD-1 trial was used based on the Intention to Treat analysis and, for the lower limit, a hazard ratio of 0.27, based on feedback from a clinician survey. The clinician survey comprised 37 clinicians (26 of consultant grade and 11 specialist registrars). The respondent group was a robust sample of clinicians treating advanced RCC in the UK (approximately 20% of the total relevant UK clinician group). The survey focused on senior clinicians who treated more than 10 Stage IV (advanced) RCC patients per year and the respondents each treated an average of 33 advanced RCC patients a year. Since Novartis has previously submitted to NICE a figure of 2,744 stage IV RCC patients in the UK, this survey covers a significant proportion (44%) of the advanced RCC UK patient population. The estimates of overall survival from this survey related to patients who had failed 1st line therapy with sunitinib and were then left untreated ie those patients that would be suitable for everolimus as per the licensed indication. The results of the clinicians' survey suggested a median OS for this patient group of 6 months (mean 6.1 months), hence a lower limit to the hazard ratio to be tested of 0.27.

However, despite being requested by NICE to investigate plausible ranges, the Appraisal Committee declined to consider the analysis based on this approach, criticising the figure used for the lower limit of the range, on the basis that "the Committee noted that these data were from a small sample of clinicians and details about the distribution of values within the data set has not been provided". For this reason, the Committee concluded that these data were likely to be "biased".

On any view, the clinicians' survey described in Novartis' response to the ACD is not "a small sample", but represents 20%, of clinicians with relevant experience in the UK. It may not be disregarded as a "small sample" and the Committee's reasons for characterising it as biased are invalid. In circumstances where the reasons given by the Appraisal Committee for its refusal to take into account the results of the Committee are incorrect, the decision is perverse.

While this is irrelevant in the context of the reasons given by the Appraisal Committee for rejecting the clinician survey, it should be noted that Novartis is able to provide further details of the survey to assist the Appraisal Committee, should this appraisal be returned to the Committee for further consideration.

Finally, it should be noted that, while the extreme assessment of uncertainty carried out by the ERG includes values that are clinically very implausible, as demonstrated by evidence to this effect provided by Novartis and recognised by NICE in its email of 16 August 2010, the Appraisal Committee has failed to take this into account in its appraisal of everolimus and this, in itself, is perverse..

2.2 The approach of the Appraisal Committee to the possibility of uncertainty in the assessment of cost effectiveness is inconsistent with that followed in other appraisals and is therefore perverse.

The Appraisal Committee's overall conclusions in relation to everolimus are set out at paragraph 4.17 of the FAD dated November 2010. The Committee referred to the deterministic ICER of £49,300 per QALY and the fact that this was associated with wide confidence intervals. In these circumstances, the Committee considered the mean probabilistic ICER of £51,700 per QALY gained. Based on a probabilistic sensitivity analysis carried by the ERG, the Committee noted that "*if the maximum acceptable amount to pay for an additional QALY gained was £50,000, the probability that everolimus was cost effective compared with best supportive care alone was only 52.7%*". The Committee's overall conclusion was, accordingly, that "*taking into account both the value of the ICERs and the uncertainty around the ICERs, the Committee concluded that it could not recommend everolimus for the second line treatment of advanced RCC as a cost effective use of NHS resources*."

However, this conclusion is inconsistent with the conclusion of the Appraisal Committee in relation to other technologies where the ICERs and the associated uncertainty were similar.

(a) Pemetrexed for the maintenance treatment of non-small cell lung cancer (TAG 190)

NICE's appraisal of pemetrexed for the maintenance treatment of non-small cell lung cancer (TAG 190) was issued in June 2010. In that case, the Appraisal Committee concluded that the most plausible ICER for pemetrexed was approximately £47,000 per QALY (paragraph 4.19) which was associated with a 57.71% probability of being cost effective at a threshold of £50,000 per QALY (paragraph 3.27). (When the Weibull extrapolation of the trial data was used, this produced an ICER of £50,673 per QALY using a probabilistic analysis, with a 49.70% probability of being cost effective at a threshold of £50,000 per QALY gained). The Committee's overall view of these data was that there was "*reasonable certainty*" that the ICER was below

£50,000 per QALY gained. Pemetrexed was accordingly recommended as a treatment option in the eligible patient population.

The ICERs accepted by the Appraisal Committee in the appraisal of pemetrexed are very similar to those calculated for everolimus and the assessments of uncertainty in the two cases are also very close. However, despite the similarity of the assessments, the overall conclusions of the Committee in these two appraisals are inconsistent. In the absence of any explanation to justify a different approach for everolimus from that followed with pemetrexed, there is a strong inference that the decision is arbitrary and therefore perverse.

(b) Trastuzumab for gastric cancer (TAG 208)

It is also relevant to compare the cost effectiveness results for everolimus to those produced in support of trastuzumab for gastric cancer (TAG 208).

The trastuzumab analysis had an ICER range of £43,200 to £52,000 per QALY gained. The range is based on a deterministic ICER of £43,200 and a probabilistic ICER of £52,000 per QALY gained. In the everolimus analysis the range was £49,300 to £50,047 with the former being the deterministic ICER and the latter the probabilistic ICER. Despite the types of ICERs considered in the range being similar, the Committee concluded that the ICER range for trastuzumab was "*within the normally acceptable CE thresholds*" and yet, when reviewing the ICER range for everolimus, the same Committee concluded that the ICER range was "*outside the normally accepted thresholds*". In addition, the trastuzumab FAD does not show evidence that the Committee even considered the probability that trastuzumab would be cost effective at a given threshold, in order to come to their decision and yet the probability of cost effectiveness was pivotal to their decision in the everolimus appraisal. There are clear inconsistencies in decision making when the two appraisals are compared and in the absence of any other compelling reasons the inconsistency in decision making is perverse.

2.3 Due to the heterogenous nature of the studies and the patient populations included in the Delea meta-analysis referred to in sections 3.6, 4.5 and 4.10 of the FAD dated November 2010, reliance of the Appraisal Committee on the ratio of PFS:OS of 1:1.4 to quantitatively justify the survival gain of 5.9 months from the ERG's analysis is perverse, as the results from this analysis are unlikely to represent so specifically the OS gain conferred by everolimus.

Results from the Delea meta-analysis, referred to in section 4.10 of the FAD dated November 2010, included 28 studies and was conducted to evaluate the suitability of Time to Disease Progression (TDP) as a surrogate measure for OS in patients with metastatic renal cell carcinoma. A number of sub-group analyses were conducted to evaluate the relationship between Progression Free Survival (PFS) and OS. Results from this analysis showed that the specific relationship between PFS and OS varies according to the studies included eg the relationship between PFS to OS is 1:1.4 for studies where there is prior therapy and 1:1.6 for studies where crossover is not an issue. However although, the Delea meta-analysis provides good supporting evidence that PFS is a reasonable proxy for OS and that the ratio of PFS:OS is likely to be 1:>1, the heterogenous nature of the studies and patient populations included in the meta-analysis means that it would be wrong to use the ratios to justify a specific,

quantitative relationship between PFS:OS for everolimus patients. Therefore reliance of the Appraisal Committee specifically on a ratio of 1:1.4 to justify the improvement in OS conferred by everolimus is perverse.

2.4 The Appraisal Committee has disregarded the available evidence for OS in patients who receive BSC

Insufficient regard has been given to the available evidence submitted in relation to OS of BSC patients following failure on 1st line sunitinib treatment.

Retrospective analyses from four large institutions, covering 220 patients who had progressed on 1st line sunitinib and did not receive any further anti-cancer treatment demonstrated that patients survived for 2 to 5 months following sunitinib treatment. In addition, results from an on-line survey, representing the views of 37 clinicians with experience of treating over 1,200 advanced RCC patients suggested that they would expect untreated patients to live for a mean of around 6 months following cessation on 1st line sunitinib. Importantly only 11% of respondents believe that these patients will survive for more than 9 months.

Furthermore a paper by Di Lorenzo *et al*. 2009, described in our ACD response is informative with regards to OS for 2nd line patients following sunitinib. The study evaluated the efficacy of sorafenib following failure on sunitinib. The median OS for these patients was 7.4 months. In many respects the patients in this study were reflective of those in the everolimus study (RECORD-1) but patients on the Di Lorenzo study could be considered as having a slightly better prognosis based on the fact patients in this study generally had better risk profiles which included better performance status and lower rates of metastatic disease in organs such as the liver, lungs and lymph nodes. Considering the fact that these patients were on active anti-tumour therapy and the patients generally had superior prognostic scores, the median OS of 7.4 months might be expected to represent a best case or superior scenario compared to the expected survival for patients who get BSC only following sunitinib.

Finally, as described in our ACD response, Liu et al. presented a poster at European CanCer Organisation/European Society for Medical Oncology (ECCO/ESMO) in September 2009, which retrospectively evaluated patients' survival following discontinuation of sunitinib or sorafenib in advanced RCC patients from routine clinical practice. The median OS results in this study for patients who only received sunitinib was 5.2 months.

Therefore the body of available evidence indicates that, if left untreated, the survival of patients eligible for everolimus is very unlikely to be more than 9 months and is more likely to be in the order of 4 to 6 months. This suggests that the ERG's modified analysis is likely to be highly pessimistic and even Novartis own estimate of cost-effectiveness is likely to be conservative. Failure to take account of these data in reaching conclusions regarding everolimus is perverse.

2.5 The estimates of mean OS associated with BSC alone, relied upon by the Appraisal Committee are inconsistent and do not reflect the referenced calculations of the ERG

The estimates of OS associated with BSC alone, relied upon by the Appraisal Committee to formulate its decision and referred to at various points in the FAD dated November 2010, are inconsistent and the conclusions based upon these estimates are therefore perverse.

- At section 4.7 of the FAD, dated November 2010, the Appraisal Committee chooses to rely upon an estimate of 10.8 months, for mean OS in patients on BSC alone, purportedly based on results produced by the ERG.

- At Table 2 of the ERG Report, dated 27 April 2010, the results suggest that mean OS for BSC alone is 10.08 months (12 x 0.84 months).

- Finally, as far as we can determine, without access to the version of the economic model modified by the ERG, the ERG's ICER of £58,316/QALY (subsequently £49,272/QALY with the revised PAS) is underpinned by an assumption that mean OS in BSC alone patients is 8.9 months.

In circumstances where OS in the BSC arm is a key driver in the cost effectiveness analysis, these differences are a matter of real concern and the inconsistencies demonstrate perversity.

2.6 Reliance of the Appraisal Committee on a mean probabilistic ICER to justify the decision not to recommend everolimus is perverse as the mean probabilistic ICER will vary from one run to another.

Section 4.17 of the FAD, dated November 2010, states in relation to the mean probabilistic ICER of £51,700/QALY gained, "It noted that this ICER was higher than those considered acceptable for end-of-life treatments to date.". However, due to the nature of probabilistic sensitivity analysis the mean probabilistic ICER will vary from run to run regardless of the number of iterations. Therefore it is perverse to justify the decision based on the results from a single run of the probabilistic sensitivity analysis, particularly as it is accepted in section 4.12 of the FAD, dated November 2010 that, by the ERG's own estimates, the probability of the ICER being less than £50k/QALY is 52.7% ie the ICER is more likely to be less than or equal to £50,000 per QALY gained.

Requested outcome

Novartis therefore requests the Appeal Panel to return this appraisal to the Appraisal Committee with the following directions:

- that the economic model as modified by the ERG should be disclosed to Novartis;

- that the Weibull curves including data employed and exaxt method used by the ERG to extrapolate OS should be disclosed or, if that is not possible, that the ERG should repeat its calculations and provide disclosure;

- that a full account of the calculation of the transition probabilities associated with the Weibull curves are provided;

- that the Appraisal Committee should consider the probabilistic sensitivity analyses carried out by Novartis using the clinically plausible range around the hazard ratio for OS; and

- that the Appraisal Committee should follow an approach to uncertainty consistent with that adopted in the appraisal of pemetrexed as maintenance therapy for non-small cell lung cancer.

Request for an oral hearing

Novartis asks that this appeal should be determined at an oral hearing.

Novartis Pharmaceuticals UK Ltd

December 2010