Dr Maggie Helliwell
Appeals Committee Chair
National Institute for Health & Clinical Excellence
MidCity Place
71 High Holborn
London
WC1V 6NA

17 January 2011

Dear Dr Helliwell,

Appeal against the Final Appraisal Determination for Everolimus for the Second Line Treatment of Advanced and/or Metastatic Renal Cell Carcinoma

Thank you for your letter dated 17 December 2010, in which you notify us of your initial views in relation to the admissibility of the points of appeal raised in our notice of appeal dated 10 December 2010. This letter sets out our response to the matters raised in your initial scrutiny letter, in advance of your final decision as to whether our points of appeal should be referred to the Appeal Panel.

Ground 1

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

We note from your letter, that you are minded to refer this point of appeal to the Appeal Panel, although you ask us to represent our arguments in a different way, as specified in your letter and to provide you with copies of documentation in which these issues were raised with NICE during the course of the appraisal.

While we recognise the reasons for the initial scrutiny of appeals and would wish to provide relevant clarification in order to assist the Appeal Panel (and, where appropriate, the Appraisal Committee) to understand the issues we are raising in this appeal, we are concerned that the request in your letter does not form part of the intended purpose of the initial scrutiny stage, as set out in NICE's Guide to the Technology Appraisal Appeal Process (the Appeal Guide). The initial scrutiny of appeals is described at Section 4.4 of the Appeal Guide, as comprising a "preliminary view of the arguability and validity of each of the points made by the Appellant" (paragraph 4.4.2). We are sure that you agree that it would be procedurally unfair for NICE to require that a valid appeal is reformulated in a particular way or to determine the way in which an appellant presents its argument.

In the context of the complex issues raised at point 1.1 of our appeal, we are concerned that these matters may not be covered adequately using the suggested approach because it is a fundamental aspect of our appeal that we have not been able to understand and investigate the economic modelling in the way we would have wished and are therefore unable to identify all the potential inconsistencies or other

problems which may be present. We have referred, in our appeal letter, to some of our reasons for wishing to investigate the modified model and those are the matters we intend to rely upon for the purposes of this appeal. However, if we were able to investigate the modified model in more detail, it is possible that there would be additional issues we would wish to raise.

Nevertheless, we have sought to present the issues consistent with your request. We hope you will understand why we believe the following summary may not be the best way of presenting our case:

- Our point of appeal relates to the fact that fairness required NICE to disclose a copy of the economic model, modified by the ERG, relied upon by the Appraisal Committee for its conclusions regarding everolimus. No reason has been provided by NICE for its failure to disclose the adjusted economic model in this appraisal.

- In relation to the economic model, Novartis had its original submission and economic model and its subsequent submissions dated 30 September 2009, 10 December 2009, 18 December 2009, 2 March 2010, 24 March 2010, 19 July 2010, and 24 September 2010; and miscellaneous emails and telephone conversations including those dated 30 June 2010, 5 July 2010, 14 July 2010, 17 August 2010, 23 November 2010.

- Novartis was provided with: the ERG report dated 30 November 2009 and the subsequent ERG reports dated 21 December 2009, 27 April 2010, 4 August 2010 and 7 October 2010; the ACD and both FADs; emails dated 1 July 2010, 1 September 2010, 6 December 2010.

- Our initial request for disclosure of the economic model, as adjusted by the ERG, resulted from our inability to understand the conclusions of the ERG as relied upon by the Appraisal Committee, as set out in our appeal letter. After the June 2010 FAD was issued, we sought and received from NICE permission to submit a revised PAS (approved by the DoH) rather than proceeding to appeal at that stage. In order to evaluate the impact of the updated PAS, we requested a copy of the ERG adjusted model, in a telephone conversation with Meindert Boysen on 30th June 2010. Instead of receiving the ERG modified model we were provided with a brief description of the ERG's modifications and the ERG's transition probabilities to input into our model so that we could replicate the ICER of £58,316k/QALY which was produced by the ERG. Although we were able to replicate the latter figure we noticed a discrepancy between the underlying estimate of OS for Best Supportive Care (BSC) quoted in the ERG's report dated 27 April 2010 and that suggested by the replicated model. Moreover, once the FAD was published further discrepancies in the OS estimates were identified, see discussion around appeal point 2.5 and Table 1 below. However, without access to the ERGs adjusted model we are unable to investigate the apparent inconsistency in the estimates of Overall Survival (OS) used at various points in the assessment. A copy of the information received is attached in Appendix 1. This point is also discussed in our appeal point 2.5.

- As a result of the fact that the adjusted model had not been disclosed, we were not able to carry out all of the analyses we would otherwise have conducted. In an email dated 16 August 2010, NICE asked us to prepare additional probabilistic sensitivity analyses. In order to facilitate the conduct of these sensitivity analyses, we requested clarification from the ERG regarding their Weibull curves, so that the uncertainty could be fully evaluated (email dated 17 August 2010). On 1 September we received a document from the ERG which suggested some of their original workings had been mislaid. In addition our other queries were not addressed adequately. Consequently we were unable to fully evaluate the appropriateness of the approach and uncertainty surrounding the ERG's extrapolation of OS using Weibull curves. A copy of the request and response is provided in Appendix 2.

- In addition, we were unable to understand or respond to certain criticisms made by the Appraisal Committee. The November FAD included a criticism of Novartis (Section 3.31), regarding the calculation of the probability that the ICER is below £50,000/QALY, stating that the ERG had identified an error in the probabilistic sensitivity analysis we had presented. The ERG 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). In an email dated 23 November 2010, we requested a copy of the ERG adjusted probabilistic sensitivity analysis. On 6 December 2010 we received an email informing us that no specific or additional alterations were made to the parameters of the model as presented by Novartis. We therefore 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 suggested by the FAD. Copies of the relevant emails are provided in Appendix 3.

- In the context of this point of appeal, the material which exists which was not provided to us, is the ERG modified economic model, clarification regarding the extrapolation of survival underpinning the modified model and adjustments to the probabilistic sensitivity analysis, relied upon by the Appraisal Committee for its conclusions regarding everolimus.

In summary therefore, in support of our case that fairness requires that the adjusted model is disclosed to consultees, we have provided examples which demonstrate (a) our inability to investigate the reliability of the modified model and the outputs reported in the FAD, (b) that we have been unable appropriately to investigate the uncertainty surrounding the ERG's extrapolation of overall survival and (c) that we have been unable to respond to criticisms regarding the probabilistic sensitivity analyses.

It is self-evident, as recognised by the Court of Appeal, in Eisai v NICE, that in circumstances where we have not been provided with the modified economic model,

it is impossible for us to provide a comprehensive list of all the matters we might raise in relation to the model, if it were to be disclosed to us and we were permitted to test its reliability. However we have sought to illustrate this point of appeal by the inclusion of examples which demonstrate how we have been prejudiced in our ability to investigate the conclusions of the Appraisal Committee in relation to the cost-effectiveness of everolimus.

Although we are surprised by your request for copies of the correspondence, as we were unaware that this forms part of the initial scrutiny stage, we enclose copies of the correspondence pertinent to the points raised above.

Finally, you indicate in your letter, that you are concerned to allow the Appraisal Committee fairly to prepare for the appeal hearing. As indicated above, we have asked NICE to disclose the modified model and have been given no explanation as to why this has not been provided. We believe fairness requires that we are provided with NICE's reasons for refusing to disclose the ERG adjusted economic model, in advance of the appeal hearing to allow us too to make our preparations.

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

In your letter, you indicate that you are minded to refer this appeal point to the Appeal Panel although, in circumstances where you suggest that the issue is closely related to the previous appeal point, you request similar information to that you asked for in relation to the point 1.1.

Appeal point 1.2 relates, like appeal point 1.1, to a lack of transparency in this appraisal, although the information to which this point of appeal refers, is different. With the same provisos as those set out in relation to appeal point 1.1, we have sought however to comply with your request:

- Our point of appeal relates to the fact that fairness required NICE to disclose information regarding the extrapolation of OS by the ERG, using the Weibull curve, including the 95% confidence intervals, which would allow Novartis to test any uncertainty in relation to the ERG's calculations.

- Clarification regarding the Weibull curves was requested in an email dated 17 August 2010. The ERGs response to this request was provided in an email dated 1.9.10.

- As indicated in our Notice of Appeal, Novartis has been unable to reproduce the ERG generated Weibull curves, used to calculate OS, using the information provided. There appears to be uncertainty with respect to the methods used by the ERG to generate the Weibull curves and the transition probabilities derived from these curves. We were unable to fully investigate this uncertainty and requests for clarification did not provide the requested explanation for the ERG's derivation of the Weibull curves and transition probabilities as these underpin the estimates of OS in the model and the incremental OS is a key driver for cost-effectiveness in the model.

- 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, Novartis could only replicate the ERG's Weibull curves if the last data point was excluded. Therefore it does not appear to be the case that all available data were used by the ERG and the Appraisal Committee acted on incorrect information when it concluded that the Weibull curve was based on all of the available data. In addition, 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 and if the Appraisal Committee has relied upon an extrapolation of OS which omits the data point from RECORD-1, this must be justified and failure to do so is unfair.

Copies of the relevant emails are provided, as requested in your letter, in Appendix 2.

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

Noted – thank you.

Ground 2

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 evidence and therefore perverse

In your letter, you say that the appeal point appears to turn on whether the only reasonable view is that the plausible range for OS, based on the results of the clinician survey and the RECORD-1 trial should be preferred over those models using the RPSFT method. You say that you are not sure how this could ever be more than a disagreement between experts. However, with respect, this does not encapsulate the focus of our appeal.

Novartis submitted two analyses for the purposes of this appraisal;
a) a scenario based on the 95% confidence intervals produced by the RPSFT methodology and
b) a scenario based on a more clinically plausible range, as requested by NICE in the email dated 16 August which stated, "We appreciate there may be additional

complexities in conducting such sensitivity analyses due to the constraints of the RPSFT method to derive the estimates of OS and as such plausible ranges should be investigated."

The issue raised in our appeal letter is that the reasons given by the Appraisal Committee for declining to consider the analysis based on the more clinically plausible range, as referred to in b) above, are inconsistent with the evidence and are therefore perverse.

In circumstances where the Appraisal Committee have rejected a piece of evidence for reasons that are perverse, its overall conclusion is clearly flawed. In this case, the reason given by the Appraisal Committee for rejecting the "clinically plausible" analysis was because it deemed the data to be from a small sample of clinicians and therefore likely to be biased. The survey included data from 20% of all clinicians treating advanced RCC in the UK and covered approximately 44% of the advanced RCC UK patient population. The survey therefore covers a very substantial proportion of relevant UK expertise in this area. To reject the data for the reasons given by the Appraisal Committee are therefore perverse.

As explained in our appeal letter, the clinician survey, conducted via doctors.net.uk, comprised anonymous market research responses from 37 senior UK clinicians who specified that they treated ten or more cases of Stage IV RCC in the last year. The reasons for conducting the survey (i.e. to inform the NICE appraisal for everolimus) and the survey sponsor (Novartis), were unknown to respondents. We have provided the raw data from the survey in Appendix 4.

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

In your letter, you refer to a recent appeal decision. Following a request to NICE to identify the appeal in question, we have been informed yesterday, 13 January 2010 that the appeal in question is that relating to the appraisal of bortezomib and thalidomide for multiple myeloma.

While therefore, we have had limited time in which to consider the background to the bortezomib appeal, we do not believe that the decision of the Appeal Panel in relation to a similar issue raised in a previous appraisal, is determinative of the issues raised in the context of our appeal. The question of whether an inconsistency in approach between appraisals amounts to perversity must depend on the particular factual situation under consideration.

By way of clarification, our appeal point relates to the fact that in two other appraisals, very similar estimates of cost-effectiveness with similar probabilities of being cost-effective at the £50k/QALY threshold were deemed to be acceptable whilst everolimus was not.

- In the case of TAG 190, pemetrexed for the maintenance treatment of non-small cell lung cancer, the most plausible ICER was deemed to be £47,000/QALY with a 57.7% probability of being cost-effective. In addition when the Weibull extrapolation of the data was used, this produced an ICER

of £50,673 with a 49.7% probability of being cost-effective at a threshold of £50k/QALY. In this case the Appraisal Committee deemed that there was "reasonable certainty". It is difficult to understand how results so similar to those of everolimus can lead to such different conclusions. This inconsistency in approach suggests arbitrariness and therefore perversity.

- Similarly for TAG 208, trastuzumab for gastric cancer, the most plausible ICER range was £43,200 to £52,000 per QALY gained. No information was provided regarding the certainty of these estimates. In the case of everolimus this range was deemed to be £49,300 to £50,047. However, in the case of trastuzumab the ICER range was deemed to be within the normally acceptable CE thresholds, whereas for everolimus, despite the upper limit being lower, the ICER range was deemed to be "outside the normally accepted thresholds".

While approaches to NHS treatment and practice may vary between different appraisals and different health technologies, the end product of the analyses, that is the ICER and the probabilities of cost effectiveness are comparable across different health technologies. For example an ICER and probability of cost effectiveness for a technology to treat obesity is directly comparable to an ICER and probability of cost effectiveness for a technology to treat cancer. It is therefore more difficult to understand why different standards and approaches should apply in different appraisals thus resulting in different decisions being made based on extremely similar estimates of cost-effectiveness. Inconsistency of approach is a recognised challenge in administrative law proceedings and we believe the concerns we have raised here justify proper consideration by the Appeal Panel of the particular issues we have raised in the context of the appraisal of everolimus.

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 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

Noted - thank you.

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

In your letter, you suggest that the Appraisal Committee considered the evidence we refer to regarding OS and rejected it and that, accordingly, our appeal is based on a simple disagreement with respect to the available evidence.

The focus of our appeal point however, is that the overwhelming thrust of all of the available evidence and views of UK treating physicians (from the survey data) and clinical experts (based on the clinical audit data) support an OS figure of 2-6 months in patients who receive best supportive care and that, accordingly, the Appraisal Committee's reliance on the ERG's extrapolation which results in a controversial and clinically implausible estimate of OS in these patients, 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

In your letter, you say that you are unable to find the ERG report dated 27 April 2010, which includes the figure of 10.08 for OS, at Table 2. The ERG's report dated 27 April 2010 is entitled "PENTAG Response to the Novartis Updated Submission (received on 24 March 2010)" and can be found at the following weblink, http://www.nice.org.uk/nicemedia/live/12044/49568/49568.pdf.

Table 2, on page 4 of the ERG's report presents OS as LYG (Life Years Gained, reported as 0.84). In order to calculate the number of months we need to multiply the LYG by 12. This means that, for BSC, OS is equal to 0.84 (row 2, column 5 of table) x 12 = 10.08 months. As you rightly point out it is the difference in OS that is important, however from Table 2, incremental OS is 0.43 (row 2, column 6 of table) ie 0.43 x 12 = 5.16 months. As highlighted in our appeal document, neither of these figures are correctly reflected in the FAD. Furthermore, when we attempted to replicate the model, there were discrepancies between those figures in the model, those reported in the ERG report dated 27 April 2010 and those quoted in the FAD. This is summarised in the table below for ease of reference.

Table 1. Summary of discrepancies in OS estimates between the ERG's modified model, the replicated model and the FAD, November 2010.

	Everolimus plus BSC Mean OS (months)	**BSC Mean OS (months)**	**Incremental OS (months)**
Replicated ERG model	14.0	8.9	5.2
ERG Report 27.4.10, page 4, Table 2 (LYG converted to months)	14.0 (1.17 LYG x 12)	10.08 (0.84 LYG x 12)	5.2 (0.43 LYG x 12)
FAD, November 2010	16.7	10.8	5.9

In circumstances where the figures for OS are crucial to the outcome of this appraisal, the inconsistency in the figures used and reported and relied upon by the Appraisal Committee raises concerns regarding the overall reliability of the Appraisal Committee's conclusions. In addition, without access to the ERG's modified model we are unable to investigate the analyses relied upon by the Committee or to explore which figures are correct.

2.6 The 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

We note the view, expressed in your letter, that it cannot be the case that use of probabilistic sensitivity analysis per se is unreasonable and your suggestion that our appeal point should not therefore proceed.

However, the basis of our appeal point is that, in the case of this appraisal, the circumstances around the application of the mean probabilistic ICER (£51,700/QALY) to justify the Appraisal Committee's decision is inappropriate. We are not suggesting that the use of probabilistic sensitivity analysis is unreasonable per se. As stated in our appeal letter, by their very nature probabilistic ICERs will vary from one run to another regardless of the number of iterations. This is because the probabilistic ICER is a mean cost effectiveness ratio of the results of the iterations and is therefore liable to outliers resulting in variability in the results each time the iteration is run. In fact probabilistic analysis is primarily intended to examine the probability that an intervention is cost-effective at a given willingness to pay threshold and not to produce a point estimate of the ICER. In this appraisal the ERG's own estimates suggest that, in 52.7% of the cases, the ICER will be below £50k/QALY i.e. the ICER is more likely to be less than or equal to £50k/QALY than not. Therefore using one of the instances when the ICER is greater than £50k/QALY to justify the decision is misleading and takes the probabilistic result out of context. In this particular appraisal, the deterministic ICER is below £50k/QALY and the probabilistic ICER will vary from run to run with a greater likelihood that the mean probabilistic ICER will be lower than £50k/QALY. This appraisal differs from most because the probabilistic ICER varies around the threshold deemed to be acceptable, although in the case of everolimus, by the ERG's estimates it is more likely to be within accepted limits.

We hope that, in this letter, we have provided sufficient information and, where appropriate, clarification, to satisfy your concerns and that you will therefore agree that our grounds of appeal should proceed to the appeal hearing. If however there are any additional areas where you require further information, we will be pleased to assist. We look forward to hearing from you as to which of the above appeal points are considered to be admissible.

Yours sincerely

Karen Jewitt
(Head of HE and OR, Oncology)