NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Health Technology Appraisal

Appeal Hearing

Sorafenib for the treatment of advanced hepatocellular carcinoma.

1) An Appeal Panel was convened on 26th February 2010 to consider an appeal against the Institute's Final Appraisal Determination (FAD) to the NHS, on sorafenib for the treatment of hepatocellular carcinoma.

2) The Appeal Panel ("the Panel") consisted of Professor Patrick Morrison (non-executive director of the Institute and chair of the Panel), Mr Andrew McKeon (non-executive director of the Institute), Mr Lester Firkins (lay representative), Dr Frank McKenna (NHS representative) and Dr Kate Lloyd (industry representative). All members stated they had no interest to declare in respect of the appeal under consideration. Mr Stephen Hocking (Beachcroft) was in attendance as a legal adviser to the Panel.

3) The Panel considered the appeal submitted by Bayer pharmaceuticals - "the Appellant". Bayer was represented at the appeal by Dr Adela Williams, Ms Nicole Farmer, Ms Sara Murray, Dr Paul Ross and Ms Noemi Muszbek.

4) In addition the following individuals involved in the appraisal were present and available to answer questions from the Panel: Professor Andrew Stevens (chair of the Appraisal Committee), Mr Meindert Boysen, Ms Rebecca Trowman, Ms Fay McCracken, Ms Francis Sutcliffe, and Dr Matt Stevenson.

5) There are three grounds on which an appeal can be lodged:

Ground 1. The Institute has failed to act fairly and in accordance with the published procedures as set out in the Institute's Guide to the Technology Appraisal Process;

Ground 2. The Institute has prepared guidance that is perverse in light of the evidence submitted;

Ground 3. The Institute has exceeded its legal powers.

6) The chair of the Appeal Committee (Dr Margaret Helliwell), in preliminary correspondence, had confirmed that the appellant had potentially valid grounds of appeal in relation to ground 1. No grounds of appeal were raised under grounds 2 or 3.

7) The Final Appraisal Determination ("the FAD") considered at this Appeal provides guidance on sorafenib for the treatment of advanced hepatocellular carcinoma.

8) Introducing the proceedings, the Panel's chair acknowledged a letter received by the institute the previous day from Professor Freemantle in relation to the economic modelling presented by Bayer and decided that the evidence put forward in that letter could be considered at the appeal. He noted that there were 4 separate appeal points and he decided to discuss each in turn.

9) The appellant was invited to make introductory remarks. Ms Farmer stated that in the first Appraisal Consultation Document (ACD) the committee had accepted that the log normal extrapolation of the data was the best fit but had changed its view in the second ACD and further changed the view in the FAD. She stated that the Committee only gave approximately 10 minutes to consider responses to consultation the meeting prior to the FAD and did not take account of clinicians' views. She stressed the innovation of sorafenib and stated that it was the first positive trial following

75 failed trials with other compounds and that it was now re-imbursed in 70 different countries reflecting the innovative nature of sorafenib. Ms Farmer also compared the approach by NICE to sunitinib which was approved for primary renal cell cancer and yet despite both sunitinib and sorafenib meeting the End of Life criteria and having similar ICERs that sorafenib was rejected.

10) For the Appraisal Committee, Professor Stevens stated that the Committee had been fully aware of the importance and innovation of the compound. He stated that the analysis of data was subject to significant uncertainty and the committee had to take this into account. He commented that although the mean increased survival in the randomised controlled trial (RCT) was 2.8 months the Committee had considered whether there may be outliers with much better prognosis. He noted that some estimates had extended the survival but also noted that there was no significant difference between placebo and sorafenib in the time to symptomatic progression and this was numerically worse with sorafenib. Professor Stevens emphasised that the Committee's role was to take an independent and fair-minded view bearing in mind the interests of all NHS patients.

11) Professor Stevens described the discussion of the Appraisal Committee regarding the difference between the ICERs for the log normal extrapolation of data and the Weibull methodology. For the Appraisal Committee, Mr Boysen stated that NICE had reluctantly accepted that the calculations made by Bayer were confidential and this made it difficult to be explicit in the documents regarding the conclusions of the Committee, in particular the results of fitting the Weibull curve could not be discussed in public. Professor Stevens stated that there were many inputs to the economic calculations that increased the ICER for sorafenib but even when using the log normal calculation and including the Patient Access Scheme proposed by Bayer, the lowest scenario value of the ICER was still very high at approximately £52,000. In view of the results of the

economic analyses, Professor Stevens commented that the Committee found the decision to be less difficult than many.

Ground 1. The Institute has failed to act fairly and in accordance with its published procedures.

12) Ms Farmer outlined the Bayer position that NICE procedures had breached appeal Ground 1 on 4 points:

a) by the Appraisal Committee failing to explain why it changed its conclusions with respect to the modelling, in the absence of new data to support doing so and by not stating the degree to which they considered evidence received during the appraisal regarding appropriate survival extrapolation methods;

b) by devoting insufficient time to considering the responses to consultation at the meeting prior to publication of the FAD;

c) by failing to place adequate weight on innovation;

d) by failing to consider the cost effectiveness of sorafenib similarly to previous compounds when applying the End of Life criteria.

Failure to explain change of conclusions on modelling

13) The chair of the Panel invited Bayer to discuss their first point in more detail. Dr Williams for Bayer explained that because of the positive outcome in the randomised controlled trial, the trial was stopped prematurely on ethical grounds. It was therefore necessary to undertake an extrapolation of the data. She argued that the Appraisal Committee had fully accepted after the first meeting that the log normal curve was the best fit of the data and this was expressed in the first ACD. The then chair of the Committee had said the issue was settled. However, she stated that after the publication of the second ACD and subsequently the FAD, the position of the Committee had inexplicably become less positive for the

log normal approach, and had cast doubt on which extrapolation methodology was most appropriate, without any other data being presented to cause this change in opinion.

14) Mr Murray gave a statistical perspective. She asserted that the log normal extrapolation was the best fit to the data and that the Weibull curve did not fit the data. She referred to there being natural history data which supported the choice of the log normal curve. In her view one could not simply "eyeball" the curves to see which seemed to fit best.

15) For the Committee, Professor Stevens denied that the Committee had changed its mind. In his view the minor changes in wording between the first and second ACD and the FAD had been to increase the clarity of the Committee's view. There was not any change in the opinion that although the log normal extrapolation gave the most optimistic view of the data, the Weibull extrapolation was also a good fit and the Committee were minded throughout that none of these analyses offered a perfect fit of the data.

16) The Panel chair quoted the letter from Professor Freemantle that the log normal was a better fit of the data. Professor Stevens remarked that at the tail of the trial data the log normal was not as good a fit as the Weibull extrapolation. Dr Murray observed that the tail of the trial data related to a smaller number of patients, and it made sense to take an approach that better fitted the data when more patients were still alive, i.e. log normal.

17) For Bayer, Dr Williams also disagreed with Professor Stevens, and stated that analysis of other data including 32 patient years of treatment had also indicated that the log normal analysis was the best fit of the data. She expressed concern that there had been a change in the chair of the Appraisal Committee during the process and this may have led to a change in view of the Committee. Professor Stevens responded by stating that this had been a very lengthy appraisal process and it was inevitable that there would be some change in personnel over time but that he had been in regular telephone contact with Professor Barnett the initial chair of

the Committee and Professor Barnett had been in full agreement with the FAD and did not consider that there had been a change in view.

18) For the Appraisal Committee, Mr Boysen described how the appraisal process had been unusual in that there had been a total of 4 meetings of the Committee as an extra meeting was convened after Bayer had submitted the Patient Access Scheme but the Committee did not produce an ACD after this meeting. However, at that meeting the committee considered a further document from the Evidence Review Group that stated in Key Issues:

> *"The PAS ICERs are very sensitive to the type of parametric fit used for survival data. The committee previously considered that both lognormal and Weibull fits to survival data were plausible. Independent survival data for advanced HCC patients (BCLC class C) tend to support this view. A key issue is then: does the range of ICERs generated by modelling these fits provide a plausible range within which lies the cost effectiveness of sorafenib for advanced HCC?" ERG p5*

19) The Committee concluded at this meeting that the reality must lie between these different analyses.

20) The Panel asked Bayer whether they felt that the log normal curve was a perfect fit or whether there was uncertainty. Ms Murray for Bayer conceded that there must be uncertainty regarding any similar analysis but that the log normal curve was the best fit of the data.

21) Mr Stephen Hocking for the Panel told Bayer that the point had been put on the grounds of an unfair consultation process, and the Panel would be asking itself whether Bayer had been able to make an intelligent response on the issues in play. Bayer were asked to comment on that test, and replied that they regard this issue as the key point in the appraisal and repeated that they did not understand the shift in the Committee's position.

22) The Panel's conclusions on this point are as follows. The Panel recognised that there was difficulty in evaluation of the data as the only RCT had to be stopped early because of the benefits of treatment from sorafenib. It was therefore necessary to undertake mathematical extrapolation of the data. The best or better approach(es) to such an extrapolation is a technical question on which reasonable people may differ. The appellant had not raised any ground two appeal on this point and it is not for the panel to substitute its own view, even if it had felt technically qualified to express one, on the detailed merits of any of the approaches in question. The issue is fairness at the time of the appraisal.

23) The Panel did not agree with Bayer that there had been a material change of position on the part of the Committee. The claim that such a change was inadequately explained therefore falls away.

24) The panel observes first, that it may be a mistake generally (and seems to have been a mistake here) to assume that even if the log normal curve had been agreed to be the "best" approach, that would be the same as it having been the "only possible" approach. That may be a possible conclusion in some cases, but it does not follow automatically. Where there is uncertainty, as Bayer acknowledged there was here, the "best" will not exclude alternatives and even approaches that are agreed not to be "best" overall may still play a role in informing an overall judgement. Professor Stevens had stated that both approaches fitted part of the data and neither approach was "fantastic overall". He stated that both fits were regarded as in a sense "extreme" and the committee took a common sense view in the middle. The panel considered that that was a sensible approach that would have been apparent to the technically informed consultee. The appellant was in a sense demanding to have something explained (i.e., a "preference" for the Weibull curve, or a "rejection" of the log normal curve) that had not been done at all. The implicit assertion that there was or should have been one curve chosen (and then explained) to the exclusion of all others was something of a straw man.

25) The Panel also felt Bayer had overstated the degree of apparent change as a matter of fact in this appraisal. ACD 1 had referred to "*considerable uncertainty*" in extrapolating the data but went on to say "*the base case lognormal extrapolation probably produced the most robust ICER for sorafenib*." ACD2 concluded "*both curve fits* [i.e. lognormal and Weibull] *were reasonable*". The FAD read "*Although the log normal curve provided a slightly better fit to the observed data, it could not be accepted as the definitive function to extrapolate beyond the data."* The panel also bore in mind the ERG report's discussion of these issues which evidenced that both curves were indeed in play, and rejected Dr Williams assertion that Bayer should be expected only to have looked at the ACD. If an appellant is asserting it cannot understand a point in an appraisal, the Panel is strongly of the view that it is incumbent on it to have read everything that bears on the point.

26) Taken in the round, and recalling that these are not legal documents designed to be subjected to minute textual analysis, the Panel did not agree that there was any material change of position to be explained, or that the Committee's position was unclear or unreasoned. The Committee began by referring to uncertainty, and saying the log normal approach "probably produced" the most robust ICER, and it ended by concluding the log normal approach "produced a slightly better fit" but "could not be accepted as the definitive function". These positions appear similar if not identical.

27) As to an explanation of why two curves were in play, the Panel considered this would be obvious to the technically informed reader from the material provided. Indeed Bayer had themselves (properly) considered a range of curve fitting approaches, before setting on the log normal approach which they considered preferable. They had submitted detailed evidence on both curves after the initial ACD. The Panel noted that that curve favoured their drug, albeit Bayer explained that that was not the reason they preferred it. The only difference between Bayer and the Committee was that the Committee kept other approaches in play. If

anything, it would have been a decision to concentrate on one curve to the exclusion of all others that would have needed to be explained. Finally Bayer had had opportunities to comment on this issue and had made good use of them. It was not apparent that there had been any lack of understanding at the time.

28) For all of these reasons the Panel rejects the appeal on this ground.

Insufficient time was allowed for consideration of the responses to consultation on ACD2

29) The Panel chair invited the appellant to outline the concerns relating to this point. For Bayer, Dr Williams stated that from feedback she had received from attendees at the meeting, the meeting that addressed the Patient Access Scheme had been too short to give adequate consideration to the points raised, and the closed session appeared to have lasted for only 10 minutes.

30) In response Professor Stevens said that in contrast, the Committee had spent more time discussing this appraisal than most, having had 4 separate meetings and a considerable amount of time had been spent by the Committee in reading the material in advance of each meeting. The appellant was not taking account of the discussions at the earlier meetings, for example, its New South Wales natural history data had been considered at an earlier meeting. At the fourth meeting Professor Stevens had given a 16-slide presentation to the Committee and the Committee had felt the decision was clear. Dr Williams for Bayer felt that insufficient consideration was given to consider new data but Mr Boysen again emphasised that the document from the Evidence Review Group had been read by the committee and included review of available data and new subgroup assessments:

"The new submission also reported numerous deterministic subgroup and sensitivity analyses using the PAS model (see Appendix 1)." ERG p7

"……..the ERG looked for published information on survival of advanced HCC patients (i.e. HCC BCLC stage C disease). As prognosis is distinctly different for Asian patients (hepatitis B regions) the ERG only sought European studies. The most relevant information came from the study by Camma et al 2008 who reported the survival of 406 consecutive HCC patients classified according to BCLC criteria." ERG p 13

31) The Panel's conclusions on this point are as follows. The Panel considered the time devoted by the Committee to the evaluation of the process. It discussed the evidence from Bayer that there was concern at the brevity of the discussion at the final meeting prior to the publication of the FAD and that at this meeting the Committee needed to consider the response to consultation, and in particular any new points presented in those responses to that ACD, and also to consider a written submission from a group of hepatologists who had argued for sorafenib to be made available. The Panel recognised that the Committee were expected to have read and digested all the written material prior to meeting.

32) The Panel noted that the Committee chair had made a 16-slide presentation at the meeting. It also recognised that this had been the fourth meeting on this topic and that the Committee had clearly spent a significant amount of time at the earlier meetings discussing the clinical response from sorafenib and the data relating to cost-effectiveness. The Panel reviewed the slides presented at the meeting and considered them to have been a fair and comprehensive review of the issues raised, and good contemporaneous evidence that the Committee had addressed the live issues at its meeting. The Panel concluded that it was a mistaken approach to focus only on the time spent in one Committee meeting, which took no account of pre-reading or of prior meetings, or of the quality or

content of the discussion at the meeting. The Panel finds that the committee had spent adequate time and taken sufficient care in considering the appraisal both generally and at the meeting in question and rejects the appeal on this ground.

The Committee failed to place adequate weight on innovation, in that its approach was inconsistent with the Institute's expectations of how innovation will be taken into account

33) The Panel Chair moved to point 3 in relation to innovation and asked the Committee to respond. Professor Stevens stated that the Committee was fully aware of the importance of innovation and explained that this was considered through the methodology of the process. In considering appraisals without End of Life criteria, the Committee were allowed to consider ICERs beyond the £20,000/QALY range and up to £30,000/QALY if the technology was innovative. Similarly for technologies considered within the End of Life criteria, the Committee could consider stretching the normal criteria in relation to ICERs considerably above this range to allow for innovation. The Committee tried to give sorafenib the benefit of the doubt but it was priced too high. There was no suggestion that there were patient benefits that were not captured through considering the treatment in this way.

34) For Bayer, Dr Ross indicated that there were likely to be a limited number of patients per annum who would be eligible for this treatment, and it was a step change compared with current treatment that was uniformly ineffective. The survival gains at present were modest, but might improve. For example, the treatment might be trialled in combination therapy. He commented that 11 hepatologists with different special interests had written a joint letter supporting sorafenib for hepatocellular carcinoma as a highly innovative new treatment. Bayer expressed concern that innovation was not expressed in the FAD but Professor Stevens considered that the FAD had fully reflected the innovative nature of the technology but the latter had been priced too high to make it cost effective.

35). The Panel considered the complaint relating to innovation and found as follows. It accepted the argument from Bayer that sorafenib was a highly innovative technology and was the only effective treatment available for advanced hepatocellular carcinoma. The Panel recognised the clinical benefits of sorafenib and that it was available in many other countries. The Panel discussed the response from the Committee chair that NICE's appraisal mechanisms addressed innovation by allowing the Committee to approve innovative technologies in the upper range of cost-effectiveness. The Panel considered whether the Committee had given due weight to the innovative aspect of sorafenib when considering cost effectiveness. The Panel recognised that the ICER for sorafenib exceeded the normal threshold of £30,000/QALY but could be considered under EOL criteria.

36) The Panel discussed the comments from the Committee chair that they had been very aware of the innovative and unique nature of the treatment. The Panel discussed the fact that there was no other treatment available for this condition and whether this had been taken into account. The Panel also discussed the relation between the cost effectiveness of the technology and the End of Life criteria. However, the Panel were persuaded that the Committee had properly considered both the EOL criteria and innovation but had not felt able to approve the technology because of the cost effectiveness data. The Panel concluded that the committee had not acted unfairly or inconsistently with the Institute's statements by not approving the technology and the appeal was rejected on this ground.

The Appraisal Committee has not explained why the magnitude of additional weight that would need to be assigned to the original QALY benefits is too great for the treatment to benefit under the EoL policy.

37) The Panel chair moved to point 4 and invited Bayer to expand the point. Dr Williams explained to the Panel how Bayer had felt the Committee's decision had not used the same criteria with sorafenib as it had done

previously with sunitinib for advance renal cell carcinoma. Dr Williams stated that the ICERs for both compounds were similar and both fulfilled the End of Life criteria, yet sunitinib had been approved and sorafenib had been rejected. This was inconsistent and unfair.

38) Professor Stevens stated that the cost effectiveness of the two compounds was not similar. He listed a number of compounds that had been approved under EOL criteria and none had had a likely ICER above £50,000 (most were substantially lower, with only sunitinib in that area) whereas the best estimate for sorafenib was above this. He stated that although there was not an exact figure specified for the upper range of the ICER when considering EOL technologies, it had previously been stated by Andrew Dillon for NICE that a multiplier of greater than 1.65 above the normal threshold was unlikely to be exceeded. For Bayer, Dr Williams stated that she had not seen a figure of 1.65, but in a NICE Board update paper summarising experience to date with implementation of the EoL criteria indicated a maximum multiplier of 1.7 that would indicate an upper range of the ICER to be £51,000.

39) The Panel asked Bayer whether even with this multiplier and the use of the Patient Access Scheme, they accepted that their own best estimates of the ICER for sorafenib exceeded this figure. Bayer accepted that this was true. In addition the Panel asked if there was a parallel to be drawn between bevacizumab plus interferon for advanced renal cell carcinoma which was rejected with the statement that:

'the lowest most plausible cost-effectiveness estimate of bevacizumab plus IFN-α of £53,800 per QALY gained' (TA178 4.3.10).

40) Bayer was aware of this ruling but considered the unique status of sorafenib should make the Committee accept this treatment.

41) For the Committee, Mr Boysen emphasised that the Committee is independent and the EOL criteria specifically leaves the decision to the judgement of the Committee. In relation to the multiplier, Professor

Stevens stated that it wasn't a concern to the Committee whether the multiplier had been 1.65 or 1.7 but all previous appraisals had been below the ICER for sorafenib and it was the view of the Committee that the cost effectiveness ratio of sorafenib was likely to be considerably greater than any previously accepted treatment. He felt the Committee could not be expected to explain in more detail why it did not make an exception for a drug with an even higher ICER.

42) The Panel considered the rejection of sorafenib in relation to previous appraisal decisions using End of Life criteria. It concluded as follows. It considered the estimated cost-effectiveness of sorafenib compared in particular with treatment of advanced renal cell carcinoma. The Panel noted that sunitinib had been approved for renal cell carcinoma and discussed Bayer's argument that this had similar cost effectiveness to sorafenib. However, the Panel noted that the estimates of cost effectiveness for sunitinib ranged below that of sorafenib and that even when the Patient Access Scheme was included, the most optimistic cost effectiveness supplied by Bayer exceeded £50,000/QALY. The Panel discussed whether the Committee had been too rigid in drawing comparison from previous technologies, but concluded that the cost effectiveness of sorafenib appeared to exceed the uplift for the End of Life criterion that had been used in previous Health Technology Appraisals. The Panel felt that it was wrong in principle for the appellant to select one treatment approved under the EoL policy, even if in some respects that treatment resembled sorafenib, and argue that that data point alone set an expectation for the policy. It may be sunitinib is an outlier. It is also wrong to look for a specific multiplier and then to seek to treat that as a new, hard edged threshold. In some cases the acceptable multiplier may be lower, in others higher. The EoL policy is a flexible departure from the Institute's usual criteria, designed to reflect a perceived higher value in life extension, and whilst it should not be applied capriciously, it should not be analysed over rigidly either.

43) The Panel agreed that as the appeal point is inconsistency the better approach was that offered by Professor Stevens, in referring to a number of past treatments approved under the EoL policy. That gives a true picture of the Committee's approach to the EoL policy, and in that light it can be seen that it is consistent with past practice not to have recommended sorafenib, on the grounds that it is insufficiently cost effective.

44) The Panel concluded that the Committee had been consistent in their utilisation of the End of Life criteria and rejected the appeal on this ground.

Conclusion and effect of the Appeal Panel's decision

45) The appeal is therefore dismissed on all grounds.

46). There is no possibility of further appeal within the Institute against this decision of the Panel. However, the decision of the Panel may be challenged by an interested party through an application to the High Court for permission to apply for judicial review. Any such application must be made promptly and in any event within three months of this decision or the issuing of the Guidance.