

15 February 2008

[REDACTED]
Ipsen Limited
190 Bath Road
Slough SL1 3XE

Level 1 City Tower
Piccadilly Plaza
Manchester
M1 4BD

Tel: 0161 870 3146
Fax: 0845 003 7785
Email: nice@nice.org.uk
www.nice.org.uk

Via email and post

Dear [REDACTED],

Single Technology Appraisal – febuxostat for the management of hyperuricaemia in patients with gout

The Evidence Review Group, the School of Health and Related Research, University of Sheffield (ScHARR), and the technical team at NICE have now had an opportunity to take a look at the submission by Ipsen. In general terms they felt that it is well presented and clear. However the ERG and the NICE technical team would like further clarification relating to the clinical and cost effectiveness data.

Both the Evidence Review Group and the technical team at NICE will be addressing these points in their reports. As there will not be any consultation on the evidence report prior to the Appraisal Committee meeting you may want to do this work and provide further discussion from your perspective at this stage. As there are a number of issues for which clarification is being requested, we have sought to list them in order of priority.

We request you to provide a written response to this letter to the Institute by end of business on Friday 29 February 2008. Two versions of this written response should be submitted; one with academic/commercial in confidence information clearly marked and one from which this information is removed.

If you present data that is not already referenced in the main body of your submission and that data is seen to be academic/commercial in confidence information, please complete the attached checklist for in confidence information.

If you have further queries on the technical issues raised in this letter than please contact Rodrigo Refois Camejo or Helen Chung. Procedural questions should be addressed to Natalie Bemrose in the first instance.

Yours sincerely

Meindert Boysen, Pharmacist MScHPPF
Associate Director - STA
Centre for Health Technology Evaluation

Encl.: checklist for in confidence information

Section A: Clarification on approach to decision problem

A1 Please discuss the clinical appropriateness of febuxostat being used in the context of a continuation rule that would allow selection of respondent patients (in terms of decrease of sUA levels and disease features) and could thereby increase the cost effectiveness of the technology. Please clarify if a switch from allopurinol to febuxostat is expected to happen in clinical practice and provide a rationale for why a sequential approach of prescribing allopurinol to all patients and then febuxostat to patients who do not respond was not considered. Please discuss in what way you would expect this to impact the analysis. Numerous studies have shown that such strategies can be more cost-effective than immediately initiating with a more efficacious but more costly treatment. Figure 5-14 shows that some patients do respond with allopurinol. The view could be taken that if such a strategy were shown to be cost-effective it would still need to be proven that febuxostat was more cost-effective than no treatment before it could be recommended for second-line therapy. It has been suggested that the following treatment strategies should be evaluated at a minimum.

1 Allopurinol (A) - Febuxostat (F) - No Treatment (NT)

2 F-A-NT

3 A – NT

4 F – NT

Where possible several switching rules should be investigated. It is currently implausible that patients would continue with medication were it to have no beneficial effect.

A2 In addition to the treatment strategies named above, please comment on why benzbromarone in combination with allopurinol was not evaluated.

A3 Please clarify if the assumption that only a minimal proportion (3%) of people receive sulphinyprazole, benzbromarone, probenecid or a combination of these with allopurinol as part of their disease management is still valid in the subgroup of patients where those technologies may be clinically more adequate, i.e. patients unresponsive or intolerant to allopurinol, and patients with renal impairment.

A4 Please provide evidence for the statement in the submission that “probenecid is not used at all in the UK” However, according to Annemans et al 2007 and prescription cost analysis (PCA) data for 2006 it is still used in England, though rare.

Section B. Clarification on effectiveness data

- B1 Please comment more on the dosage of allopurinol within the pivotal head-to-head RCTs - in particular on why allopurinol doses were not escalated in non-responders and on the percentage of patients who had subtherapeutic doses of 100mg. The submission suggests that “allopurinol has limited efficacy in lowering sUA”. This statement is supported by evidence from the FACT trial which used fixed dose (300mg/d) allopurinol. However, in practice allopurinol is titrated in increments of 50 to 100mg, up to 900mg/d, to achieve the targeted sUA level. Is there other trial evidence to support this statement?
- B2 Please discuss how the results for allopurinol in the head to head trials compare with those from previous allopurinol trials and whether this is likely to impact on the results. It is noted that the search strategies did not include the word allopurinol. Whilst this may have been due to the large head-to-head trials that were known about, if previous allopurinol trials had contradictory evidence a mixed treatment comparison model could be needed.
- B3 Please provide a rationale for the presenting the pooled analysis of individual patient data from the FACT and APEX trial and discuss the methodological limitations of doing so. Could you provide a meta-analysis of the safety and efficacy evidence (as noted in guidance section 5.5) using individual patient level data or aggregate data for all outcomes of interest and not just significant results? It is not explicitly clear why a meta-analysis was not undertaken. A pooled analysis primarily focuses on treatment groups rather than on studies and this approach does not consider the validity of the comparisons; thus it is subject to confusion bias (also known as Simpsons paradox in probability). A further consideration which supports meta-analytical techniques is that the study design, population, and follow-up are subtly different. Also, please comment on the validity of pooling end of follow up data (i.e. endpoint at final visit) from two different trials with distinct follow up times (28 or 52 weeks) and provide a clinical rationale to justify that approach.
- B4 Please give information on the number and time profile of gout flares experienced in the febuxostat and allopurinol groups in both phase III clinical trials. Also, could you provide more details on severity of gout flares and how this was analysed and its implications (Section 5.3.4 Outcomes)?
- B5 Please explain what the implications of the high rate of flares with febuxostat treatment are, particularly within the first year. Please provide appropriate data in table format; section 5.4.5 (secondary endpoints: proportion of patients requiring treatment for gout flare) has been poorly reported with important omissions. It appears that limited/selective data have been reported, particularly on the number of patients that experience gout flares from day 1 to end of study. Only data for the last 4 weeks of the FACT and APEX trials have been reported (Figure 5-6 does not report n/N at each time point). It is

noteworthy, that in the FACT trial, nearly 70% of patients in the febuxostat groups experienced gout flares after discontinuing prophylaxis i.e. between week 9 and 52 (no data reported for APEX trial from week 1 to 24).

- B6 Please provide further detail of discontinuation rates for the APEX trial? It is unclear why this was not reported in the submission. Although the rates of premature discontinuation have been reported (Table 5-7, Table 5-8 and 5-19), it is not clear if these are significant for both febuxostat groups vs. allopurinol. In the FACT trial, the rates of discontinuation were significantly higher in both the 80mg febuxostat group and the 120mg febuxostat group than allopurinol group ($p=0.04$ and $p=0.002$, respectively). Please also comment on the paper by Pohar and Murphy in [Issues Emerg Health Technol.](#) 2006 Aug;(87):1-4 that the discontinuation rate on febuxostat was larger than on allopurinol.
- B7 Please clarify whether there is evidence for no dose adjustment in the elderly? Reference 9 appears to be based on clinical opinion rather than trial evidence.
- B8 Please discuss the following: Section 4.5 (p 25) emphasises that under-treatment is to do with clinical practice and only partly to do with drug effectiveness. How will febuxostat prescription ensure that the other drawbacks of current practice are allayed?
- B9 Are serum uric acid target levels, as recommended by EULAR and BSR, based on consensus or on empirical evidence?
- B10 What is the strength of evidence on when to start urate-lowering drug therapy for chronic gout i.e. two/three/four acute attacks per year?
- B11 Please comment on the clinical perspective of not providing prophylaxis after flares. If there are different numbers of flare on each treatment arm then omitting these costs in the economic model will have an impact on the results.
- B12 please provide a QUORUM diagram? Section 5.2.3. states “a flow diagram of the number of studies included and excluded at each stage should be provided...as per the QUORUM statement flow diagram.”
- B13 Please provide clarification on referencing of trials?; there appears to be some confusion (Methods, page 38).
- B14 Please provide justification and explanation for the approach covered in Section 5.3.5 statistical analysis? Ideally in an intention to treat analysis participants should be analysed in the groups to which they were randomised regardless of which (or how much) treatment they actually received, and regardless of other protocol irregularities, such as ineligibility.
- B15 Please clarify what are the differences between measurements at final visit (when were these taken?) and at 52 weeks in the FACT trial. Similarly, in the APEX trials what are the differences between measurements at final visit and 28 weeks. Which one of these is the most appropriate measure?

- B16 Please provide data for number of tophi with febuxostat 80mg/d in the APEX trial (p.65)?
- B17 Please provide clarification on the following: the text for Table 5-17 suggest the results are from last 3 months but the corresponding table presents final visit results – clarification and consistency in reporting results is needed, e.g. findings cannot be compared with Table 5-16.
- B18 Please provide data for each trial (APEX and FACT) as baseline, final and mean % change on the secondary endpoint analysis of tophi resolution section 5.5.6. Data in the text and table 5-25 appear to be from the FACT study only and percentage changes of primary tophi in Table 5-25 are reported as medians and not means as suggested in title.
- B19 Please provide data for each study (and treatment group) as well as further clarity on the degree, duration and severity of adverse events (including definitions). In particular you should provide further data regarding the reason for deaths (including dose) by each study and treatment group.
- B20 Please explain the significance of the following results and provide a full breakdown of withdrawals due to (type and severity) adverse events for each study and treatment groups. In the FACT trial, the most common adverse event leading to withdrawal was abnormal liver-function test results which accounted for withdrawal of 5 patients receiving febuxostat 80mg/d, 7 receiving febuxostat 120mg/d and 1 receiving allopurinol ($p=0.04$ for febuxostat 120mg vs. allopurinol). In addition, more patients receiving febuxostat (80 or 120mg/d) discontinued the study because of rashes. It is not clear, why this has not been reported in the STA submission. It would also be helpful if you provide further details on the other reasons for withdrawal i.e. why are there large (significant?) differences between and within febuxostat groups and allopurinol in personal reason, other, protocol violation and therapeutic failures (how defined)?
- B21 The last row of Table 5-15 is confusing. One suggestion would be to split it into two rows one versus placebo and one versus allopurinol. At present it does not show if allopurinol is significantly better than placebo.

Section C. Clarification on cost-effectiveness data

- C1 Please provide the rationale that each sUA band is associated with a certain utility value (independently of the number of flares). Provide evidence that the utility in a person with sUA < 6 with a flare would be greater than a person with sUA > 10 who does not suffer a flare. This data may be biased as a high sUA could be caused by other diseases, such as hypertension which could be the reason for lower utility levels. Additionally clarify how those values were retrieved and how were they controlled for the number of gout flares in order that the disutility of gout flares was not double counted. If these were successfully controlled for, discuss whether subtracting the same disutility for a flare from all sUA levels is correct, as those in worse sUA states have a lower utility decrease when moving to the assumed 0.5 utility during the period of the flare. Justify the assumption regarding the disutility associated with flares. Please provide the full IMS Report reference 12 as an appendix to your submission.
- C2 Please provide a clinical rationale for the assumption in the model that the sUA levels achieved after the first 12 months of treatment will be maintained throughout the second year of the analysis. Please discuss to what extent these results compare with the data from the follow up extension longer term studies and justify why those have not been used in the extrapolation exercise. It is noted that the proportion of patients amongst sUA levels changes in subsequent years. Please provide clarification on the data and sources for these assumptions and how these compare with epidemiological studies of sUA levels in patients treated for gout. Note that reference 37, Roddy et al, report that in treated patients 77% had sUA < 6.
- C3 Please provide as much detail as possible on the methodologies used in the logistic regressions performed to estimate the relationship between sUA and utility; the relationship between sUA and odds of experiencing gout flares; and the methods used to calculate the monthly average of flares experienced per patient.
- C4 It is important to clarify the final product price to be submitted in the UK. The price of febuxostat used in the analysis appears to be uncertain. Note that the currency conversion rate has significantly varied since October 2007). The cost of drugs should not be in the probabilistic sensitivity analyses as those are fixed by a list price. Setting cells u24 and u25 in the PSA_Febuxostat_pooled worksheet would appear to correct this problem. Note that the cost of allopurinol 300mg is 0.0643 pence per day, not 0.065 as used in the model
- C5 Please check the following the formula in O24 in PSA_Febuxostat_pooled which is incorrect.
- C6 Please check the methodology for calculating median ICERs; which is incorrect. This method will only work if the ranks of each cost and QALY pair are identical. The only way to do it is to calculate the ICERs and find the 50th percentile.

- C7 Please confirm that all reported ICERs are correct. The 95% confidence intervals in Table 6-23 are incorrect. At month 24, over 5% of incremental QALYs are negative, and no costs are negative, thus at least 5% of the ICERs must be dominated. The upper 95% CI for the ICER is however reported as confidential data.
- C8 Please provide clarification on Table 6.4 in the context of the model. The mean data for allopurinol is stated in the table as 0.2016, but is in the model as 0.2108. The distribution from the raw data is also unlikely to be normal. Clarify that the febuxostat data is also correct within the model. Comment on the patients with extremely high number of flares (>10 flares).
- C9 Please clarify the following regarding the algorithm that has been used to calculate the reduction in the initial flare rate with prophylactic treatment. It is stated that it was based on two 6-month controlled studies of colchicine versus placebo and was then applied to the flare rate retrieved from the clinical trials data. Were these two studies meta-analysed to give the 78% estimate in flare reduction? Please discuss the comparability between the population in that study and the population in the referred clinical trials. It is also stated on the clinical trial protocols that patients received prophylactic treatment for the first 8 weeks during the studies. Please clarify whether the application of such flare reduction algorithm represents an overestimation of the prophylactic effect in reducing initial flares due to 8 week prophylaxis period.
- C10 Please clarify whether the number of flares was an endpoint in the clinical trials and justify why those have not been used in the model, but instead a study was commissioned to investigate the relationship between sUA and number of flares. Also, please clarify whether the logistic regression used to estimate that relationship applied any form of control for other gout manifestations as presence of tophi. Give further details on the rationale for only applying the logistic regression after 3 months of treatment, and ignoring this relationship within the initial 12 week period. Comment on how well the relationships established between sUA and flares fit the data provided in the RCTs.
- C11 Please justify on clinical grounds the economic model structural assumption that the probability of experiencing a flare depends on sUA levels but the number of flares experienced do not. Also, taking in consideration that the experience of a flare does not have impact on the patient's path on the model, please clarify why an approach where the average number of flares per patient in each sUA level was not used.
- C12 Please clarify why the prophylaxis success rate was not included as a parameter in the probabilistic sensitivity analyses. Also why was the proportion of patients in states other than sUA < 6 not included, and why was the increased rate of flares in different sUA bands not included?

- C13 Please justify why the economic model presented assumes only 3 months of initial flare triggering from urate lowering therapy. Could you provide sensitivity analysis on the extension of the first phase of the model up to 6 months, noting that BSR Guidelines state that initial prophylactic treatment can go up to 6 months.
- C14 Please state if compliance with febuxostat is expected to be similar to allopurinol and to what extent will this impact the results of the cost effectiveness analysis. It is suggested that attrition rates for patients receiving allopurinol are relatively high...
- C15 Please comment on the strength of evidence (e.g. meta-analysis of RCTs, RCTs, cohort etc) for linking serum uric acid levels (surrogate outcome) and clinical outcomes (e.g. gout recurrence, reduction in tophi)?
- C16 Please provide confirmation on the utility values used; co-morbidities appear to be incorrect and not those reported in the Anneman et al paper for the UK.
- C17 Please clarify the methodology for obtaining the CEACs. Traditionally this would not have more than 3 dp as only 1,000 simulations were provided, whereas the figure reported for being less than £20k at 24 months is 0.6234.
- C18 Please provide a rationale for not undertaking any subgroup analyses when in this disease's context specific subgroups may imply different disease management (e.g. patients with renal impairment) and eventually require different clinical endpoints (e.g. patients with tophi requiring a quicker reduction in sUA levels) from the overall population. It is stated that not enough evidence is available to differentiate specific subgroups of patients in terms of clinical effectiveness. It is also suggested that a logistic regression showed that the treatment population is fairly homogeneous when defined by age, gender and disease severity but not comment is made with regards to people with renal impairment or people intolerant or irresponsive to allopurinol. Please consider that individual patient-level data is available and that overall population results may mask higher/lower clinical and cost effectiveness in specific subgroups.
- C19 Please comment on the likely affect of incorporating this compliance within the model. It is noted that more patients dropped out of trials with liver dysfunction on febuxostat than allopurinol.
- C20 Please clarify in a clinical context why within the model the dosing of febuxostat is allowed to be altered (letting patients move from 80mg to 120mg) but the dose of allopurinol remains constant. Given that costs for assessing maintenance treatment of gout patients are included the data to allow allopurinol titration should be readily available.
- C21 Whilst the Beta distributions are handled correctly in the PSA, they become implausible in the univariate analyses, where proportions exceed 1. Confirm that this analyses is not included in the report.

- C22 No correlations are included in the PSA. For example the mean numbers of flares across time (T19:T23) in the PSA_Febuxostat_pooled worksheet are likely to be correlated. The present methodology will tend towards average results, as the 5 are independently sampled. Comment on the effect on the results were correlations to be included between this set and amongst other sets of data.
- C23 Clarify why the theophylline testing that is commented upon in the SPC when initiating febuxostat is ignored within the model.
- C24 Clarify Table 6.6 How does the probability of flares relate to the mean monthly number of flares.
- C25 Clarify Table 6.13 and 6.14 Presumably a total cost has been divided equally. Comment within the text
- C26 Table 6-10 – the confidence intervals are incorrect see Table 6-4.
- C27 Presumably the € within the model should in fact be £?

Section D. Additional issues

- D1 Why are the inclusion criteria, namely the duration of study, different between the clinical (no restriction) and cost (minimum 12 weeks) section?
- D2 Reference 9 seems to be incorrect linking to a different paper.
- D3 Table on page 10 is in conflict with the text. $8 * 59.5 = 474$