National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Abiraterone acetate for the treatment of metastatic hormone relapsed prostate cancer not previously treated with chemotherapy

You are asked to check the ERG report from Kleijnen Systematic Reviews to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm** on **27 March 2014** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

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

Issue 1 Whether abiraterone acetate meets end of life criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page10 According to the manufacturer abiraterone meets the criteria for appraisal of end of life medicines. However, looking at the COU-AA-302 trial data it is unlikely that life expectancy in this patient group will be less than 24 months…as far as the ERG is aware the "short life expectancy, normally less than 24 months" is based on the normal treatment options available for these patients without the intervention under assessment. Page 111 Regarding the first criterion, Figure 6 in the MS (see Figure 7.1 below) shows that after 24 months, approximately 63% in the control group are still alive; and that the median survival is 30.1 (95% CI: 27.3 to 34.1) months. Therefore, it is unlikely that life expectancy in this patient group will be less than 24 months.	Janssen believes that the statement that 'it is unlikely that life expectancy in this patient group will be less than 24 months' is not factually correct and requests that this statement is removed and is replaced by a more balanced interpretation of the available evidence on life expectancy	As stated in our submission, the European Association of Urology (EAU) guidelines state that overall survival for 'hormone refractory' patients range from 9-27 months; the higher estimates of 18-27 months reflect those with only minimal metastases. Furthermore, a systematic literature review by Kirby *et al* (2011) identified 12 publications that included survival estimates in patients with mCRPC; median survival was found to range between 9-30 months. A pooled, sample-weighted median survival of 14 months was calculated from the data. Kirby *et al* explicitly state that although estimates ranged from 9 to 30 months, "for those with metastatic disease, this was reduced to 9-13 months, confirming the impact of metastases on survival". In addition, radium-223 is currently undergoing the NICE TA process, and the ERG (Aberdeen; ERG report released on 24 March 2014) who have reviewed the manufacturers' submission state that end of life criteria are met for this patient population, which is	We have described the results from the COU-AA-302 trial and drawn our conclusions based on those data. There are no factual errors in our description of the data. Therefore, no changes have been made to the report.

Page 112 According to the manufacturer, patients in the trial are likely to have gone on to receive other clinical trial technologies post-docetaxel and therefore the survival observed for these patients is probably not reflective of the average mCRPC patient in the UK. However, as far as the ERG is aware the "short life expectancy, normally less than 24 months" is based on the normal treatment options available for these patients without the intervention under assessment.		highly similar to the patient population relevant to our submission. Lastly, we maintain that the patients in the control arm of the COU-AA-302 trial were eligible to receive experimental treatments post-docetaxel, and therefore the median survival of 30.1 months observed is not reflective of patients in UK clinical practice. Taking into account all of the above, Janssen asserts that it is highly unlikely that the patient population relevant to this submission will have a life expectancy of more than 24 months.	

Issue 2 Whether abiraterone acetate meets end of life criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page111 Regarding the second criterion, the manufacturer provides median survival estimates, but not mean survival in the MS.	Janssen requests that this statement is removed.	Janssen believes that this statement is not relevant to the consideration of whether AAP meets the end of life criterion, compared to current NHS treatment. We conducted a focused review of all previous HTA submissions that have appealed to NICE's end of life supplementary advice since its introduction in 2009. Of those considered by NICE to have met the end-of-life criteria of	For the assessments of the End-of-life criteria, the Appraisal Committee usually wants to know mean and median survival for each arm and mean and median survival gain of the intervention over BSC or placebo. No changes have been made to the report.

		an additional 3 months survival, *median* overall survival estimates, not *mean were accepted in almost all appraisals*. In the COU-AA-302 trial, AAP was associated with an increase in median OS of 5.2 months for AAP when compared with BSC (PP) in the post-ADT pre-chemotherapy setting. In our economic model, AAP was associated with an increase in median OS of 7.44 months when compared with BSC (PP). Furthermore, 45% of patients from the COU-AA-302 trial were still alive at the 3rd interim analysis; therefore, by default, the calculation and provision of a *mean* overall survival estimate based on a 55% data cut would be inappropriate. This is often the case in oncology trials where the mean OS is not yet reached by the study cut-off date.	

Issue 3 Mean overall survival (OS)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 64 Janssen was unable to provide mean survival for both groups or mean survival gain, despite explicit questions in the	Janssen does not believe that this is a fair interpretation of the situation and requests that further context is provided to the Committee to ensure they understand the reasons for this. *Mean survival time in the COU-AAA-302 trial is essentially meaningless: as the majority of*	This statement needs to be contextualised with the justification as to why we could not provide the mean survival data.	The ERG believe it is for the Appraisal Committee to decide whether data are relevant or not. Fact is that Janssen was unable to provide these data. Therefore, no changes have

clarification letter.	*patients are still alive at end of follow-up and the trial is still on-going (note that 63.4% of patients in the AAP arm and 56.8% of patients in the PP arm were censored at the 3rd interim analysis, the latest data cut available), the calculated mean number of days of survival is by definition an under-estimation of the actual survival (and more so in the AAP arm compared to the PP arm, as more AAP-treated patients were still alive at data-cut). In order to calculate a mean, one requires the full data set. Consequently, in oncology clinical trials of this nature, the median overall survival is used as the measure of efficacy, as opposed to the mean.*		been made to the report.

Issue 4 Choice of model structure

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 14, 102 A less elaborate Markov model (using more health states than the regular three-state model) would have been as appropriate. Page 15, 76, 102, 106, 126 The ERG believes that it would have been possible to use a more transparent model with less phases and equations.	Janssen does not believe these statements are correct, and would request that they be amended to reflect a more accurate reflection of the choice of model structure.	Using a Markov model would have been less transparent than the DES model. The major limitation of a Markov model is that it has no memory. The probability of transitioning from one state to the next is independent to the previous. Past patient experience (e.g. duration of treatment) in the mCRPC pathway is essential in determining what happens next (e.g. treatment discontinuation/ switch). To minimize this limitation,	According to the ERG, the model provided by the manufacturer could have been more transparent. ERG report Page 76 "While the manufacturer considers the model presented as "a simple discrete event simulation (DES) model" (pg. 18, MS1) the ERG does not believe that a DES model, simulating individual patients

		numerous health states and tunnel health states would have to be created, decreasing the model transparency. One example of this is TA216 (Bendamustine for CLL), where a Markov model with 39 health states was required to track two lines of treatments. Transforming time-to-event data (as we have for the COU-AA-302 trial) to generate transition probabilities is not an efficient use of the data, as these are time dependant probabilities. Using a DES model allows one to model this more accurately. Moreover, given that we have complete patient level data for the treatment under appraisal (AAP) vs the relevant comparator (BSC, PP), a DES model allows us to use all of our data to track each individual patient, and to model the trial population with a high degree of accuracy.	using 17 prediction equations would have been the simplest and most transparent approach. The ERG believes instead that it would have been possible to use a more transparent model, for instance a Markov model consisting of health states according to the treatment phases included in the current model and a sufficiently short cycle time. This model would also allow reflection of the clinical pathways in the UK and to produce results for subgroups with varying baseline characteristics. … Transparency is a key aspect of modelling and in this specific case a more transparent model would be more convenient for an external reviewer to assess face validity and internal validity of the model." The ERG acknowledges the memoryless property of a Markov model. This however can be relaxed using tunnel states (as described by Janssen). Moreover, specialized software such as TreeAge can be used if the number of health state becomes too large (see for

			instance TA 176). This is not a factual error. Therefore, no changes have been made to the report.

Issue 5 rPFS as an established endpoint

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 27 rPFS is not an established end point in metastatic prostate cancer. This is the first time that rPFS has been used as a co-primary end point in this setting and it is not clear if rPFS superiority translates into clinical benefit.	Janssen believes that this statement is factually inaccurate and requests that this sentence is removed.	Radiographic progression-free survival (rPFS) is an established end point in metastatic prostate cancer. It has been used as an outcome in numerous studies of the condition, for example the TROPIC registration study of cabazitaxel (de Bono et al, 2011), the IMPACT registration study of sipuleucel-T (Kantoff et al, 2010), the AFFIRM registration study of enzalutamide post-docetaxel (Scher et al, 2012) and most recently the PREVAIL registration study of enzalutamide pre-docetaxel (Beer et al, 2014). In all of these studies, the range of efficacy results demonstrated including rPFS led to the licensing of the respective drugs with the regulatory authorities. A retrospective analysis of the results of the COU-AA-302 study has demonstrated a 'robust association between rPFS and overall survival' (Ryan et al, 2012) suggesting that rPFS superiority does in fact translate into clinical benefit.	As far as the ERG is aware this is the first time that rPFS has been used as a co-primary end point in this setting. The TROPIC study (De Bono et al. 2010) used OS as the primary outcome and PFS (defined as "from the date of randomisation to whichever event occurred first: prostate-specific antigen progression, radiological progression, symptomatic progression, or death) as a secondary outcome. AFFIRM (Scher et al, 2012) used OS as the primary outcome and secondary end points included measures of response (in the PSA level, in soft tissue, and in the quality-of-life score) and measures of progression (time to PSA progression, radiographic progression-free survival, and time to the first skeletal-related event). We do not have access to Beer et al. However, this study was

			published after COU-AA-302. Therefore, no changes have been made to the report.

Issue 6 Validity of subgroup analyses

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 101 The manufacturer states that the HRs estimated for the prespecified subgroups were consistent with MS Figures 7 and 8. However, the HRs and their 95% confidence intervals were not specified. Moreover it was unclear how the estimated HRs were exactly incorporated in the DES model. Thus the validity of the subgroup analyses could not be established.	Janssen requests that the statement 'Thus the validity of the subgroup analyses could not be established' be removed.	As the model tracks patients through multiple phases of treatment until death, HRs of OS were not directly used in the model (for both overall population and subgroups). Treatment effects on survival at each phase of treatment are included in the prediction equations. As the model directly samples patients from the COU-AA-302 trial with patient characteristics, subgroup analysis was conducted by including patients with corresponding characteristics. For example, for the BPI 0 or 1 group, patients whose BPI was greater than 1 at baseline from the trial were excluded from the sample. BPI and ECOG are both significant predictors in the predictive equations for patient survival. Their coefficients in these predictive equations are contributing to the predictions of subgroup analysis. The interaction term BPI*AAP is a statistically significant coefficient, which implies that the AAP arm has a different survival impact amongst different BPI subgroups.	The description of the methods used for the subgroup analyses in the MS was brief. As a result the ERG was unable to verify the validity of these analyses based on the information provided in the MS. This is not a factual error. Therefore, no changes have been made to the report.

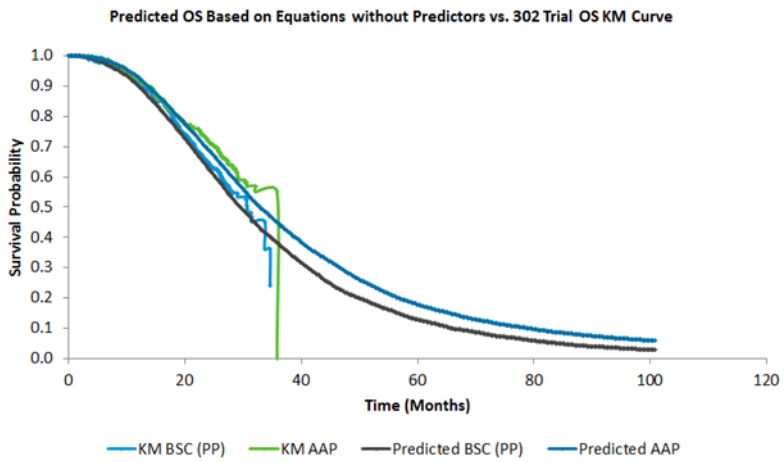
Issue 7 Use of ITT population in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 14, 79, 80 It would be more plausible to use the ITT population (despite missing baseline information) instead of the selected subset of patients.	Janssen requests such statements be revised to the following: "Whilst the ERG believes it would be more plausible to use the ITT population for the statistical analyses, it recognises that baseline characteristics are fundamental in generating risk equations in a DES model and a revised approach may be necessary."	Utilising the ITT population would not be more plausible since risk equations in a DES model cannot be generated without baseline characteristics. In addition, the analysable patient group has similar survival to the ITT population (the difference is not statistically significant).	The ERG disagrees that risk equations in a DES model cannot be generated without baseline information. Covariates should only be included in equations if plausible and if it provides more valid estimates. In this situation, there was a trade-off between (presumed) increase in validity due to the incorporation of covariates and the (presumed) loss of validity due to the exclusion of **17%** of the ITT population due to missing baseline information. This is not a factual error. Therefore, no changes have been made to the report.

Issue 8 Bias in the model in favour of AAP

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 14, 79, 85 By excluding patients with missing baseline information, bias in favour of AAP has been introduced in the model for both TTD and OS…. indicating an increased median TTD and an increased median OS by 1.05 and 0.27 months respectively favouring AAP.	Janssen request such statements be removed based on factual inaccuracy.	Using the analysable data set did not introduce bias in favour of AAP. The median TTD for AA is estimated to be 1 month longer in the analysable data set, while the median survival is estimated to be only 0.27 months (8.4 days) longer. The additional 8.4 days of survival is of insufficient length to outweigh the substantial addition of 1-month drug cost of AA. Hence the incremental impact of using the analysable data leads to an overall increase in the ICER. This does not bias in favour of AAP.	The manufacturer argues that the additional median survival of 8.4 days is of insufficient length to outweigh the substantial addition of 1-month drug cost of AA. However, in the ERG report it is stated that 'bias in favour of AAP has been introduced in the model for both TTD and OS'. There is no mentioning of bias in favour of AAP in terms of cost-effectiveness. This is not a factual error. Therefore, no changes have been made to the report.

Issue 9 Prediction equations

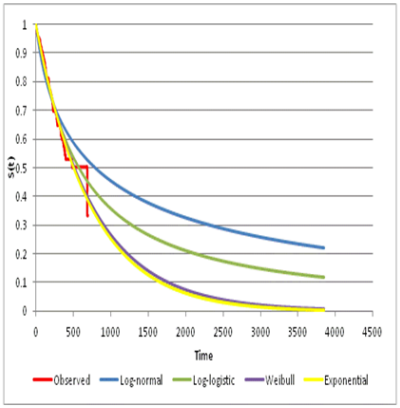
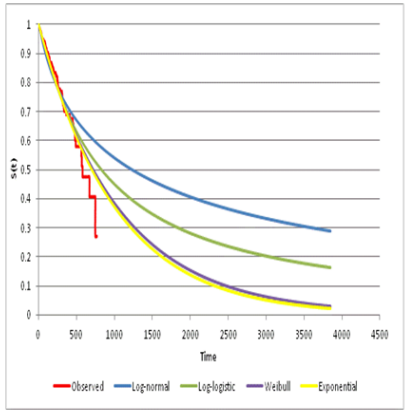
Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16, 85, 106 Prediction equations…were based on the ITT population and consistently including treatment as the only covariate for the ERG base case.	Janssen requests that such statements be revised to reflect what is feasible in a DES model.	It is not possible to input unknown baseline characteristics in a DES model and use treatment as the only covariate in the ERG base case. Using the ITT population with treatment as the only predictor cannot link a patient's experience, and so results in an underestimate of OS compared with the COU-AA-302 trial OS KM curve.  	This is not a factual error. Therefore, no changes have been made to the report. See our response on issue 7.

Issue 10 Incorporation of possibility of dying during all phases of the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 76, 82 It would be more plausible to incorporate the possibility of dying during all phases in the DES model.	Janssen requests such statements be removed from the report.	This is an inaccurate comparison between a DES approach and a Markov approach. In a Markov model, the probability of death is always considered for each non-death health states at each model cycle. However, a DES model is driven by a chronological sequence of event occurrences. When patients die whilst on treatment, their death and treatment discontinuation happen simultaneously (ie, a patient cannot be treated once they have died).	In the DES model for the time on docetaxel two equations are used: time until treatment discontinuation and time until death. The event that comes first is applied. For other treatment phases (i.e. pre- and post-docetaxel active treatment) only time until treatment discontinuation is modelled. The ERG argued that it would be more plausible to incorporate the possibility of dying during all phases in the model. This is not a factual error. Therefore, no changes have been made to the report. Besides this, in the clarification phase, the manufacturer responded to this same issue by stating that “Since so few deaths were recorded while on treatment, it was not possible to create a competing risk equation for death versus treatment stop, instead these deaths were considered in the derivation of survival post AAP/BSC (PP) treatment discontinuation (with time of death post AAP/BSC (PP) discontinuation set to 1 day for these 5 patients).

			Similarly, while patients were treated with post-docetaxel active treatment, … The limited number of death events while on post-docetaxel treatment was not enough to power the statistical analysis to create a competing risk equation for death versus treatment discontinuation. As in the pre-docetaxel setting, prediction equations were developed to predict the time to death from the time of post-docetaxel active treatment discontinuation. The deaths observed while on post-docetaxel treatment were considered by setting the time of death to day 1 after discontinuation of post-docetaxel active treatment.”

Issue 11 Prediction equations estimated separately by arm

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 14, 82, 83 The prediction equation for "Time from AAP/BSC (PP) end to death" was, unlike all other prediction equations, estimated separately by arm… the manufacturer could not provide a convincing reason for using this procedure.	Janssen request this be reworded as follows: "The prediction equation for "Time from AAP/BSC (PP) end to death" was the only prediction equation which needed to be estimated separately by arm since utilising a single equation with treatment as a predictor was inappropriate."	Selecting a single equation with treatment as covariate or separate equations for each prediction equation was based upon visual inspection. For time from AAP/BSC (PP) end to death, a single equation with treatment as a predictor did not seem to provide a good fit, though single equations provide good fit for other prediction equations. See comparison of figures below: AAP Arm  BSC (PP) Arm  	The ERG argued that (ERG report Page 83): "The prediction equation for "Time from AAP/BSC (PP) end to death" was, unlike all other prediction equations, estimated separately by arm (see MS1 Table 148). After receiving the observed KM and cumulative hazard functions (response to clarification question B5) which were used by the manufacturer to decide whether treatment should be included as a predictor in the model or stratified models should be estimated for each predictor, it was unclear to the ERG why the "Time from AAP/BSC (PP) end to death" prediction equation stands out and a stratified model was used for this prediction equation. Therefore, to be consistent

			with other prediction equations, a non-stratified equation would be preferred by the ERG.“ This is not a factual error. Therefore, no changes have been made to the report.

Issue 12 Ordering of steps

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 83 The order of the steps to estimate and select the prediction equations should be changed.	Janssen requests this statement is removed from the report.	This suggestion is not practical as a parametric distribution has to be specified when testing the significance of the covariates. Otherwise, the covariates have to be tested in all parametric distributions, with each distribution having a set of selected covariates, which is extremely inefficient and unnecessary.	We agree the approach we suggested is more labour some. However, textbooks on survival analysis suggest that adding covariates may alter the selecting of an appropriate parametric model. See for instance Chapter 10 in Statistical Methods for Survival Data Analysis by Elisa T. Lee & John Wenyu Wang. Wiley, 2013. This is not a factual error. Therefore, no changes have been made to the report.

Issue 13 Censoring

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16, 79 The remaining 15% of the patients, who did not follow the specified treatment pathways… were censored at the time they deviated from the pathway.	Janssen requests this statement is removed from the report.	This is a factual misrepresentation. Patients who received further active treatment after post-docetaxel active treatment are not censored. Their time spent with these treatments was included in the time from post-docetaxel treatment end to death.	The paragraph cited in the description of the problem refers to a particular quote in the ERG report coming directly from the MS (Page 271): “The majority of patients (85%) followed the pathway of treatments specified in the model; however, the remaining 15% patients who did not follow the treatment pathway (e.g. placebo patients started AAP before docetaxel) were censored at the time they deviated from the pathway.” This is not a factual inaccuracy. Therefore, no changes have been made to the report.

Issue 14 Comparison of post-docetaxel survival with TA259

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 15, 100 Table 5.16: Comparison of post-docetaxel survival with TA259	Janssen request these analyses be correctly run.	TA259 data has been applied wrongly by the ERG in this analysis. TA259 data should be applied to both "time from post-docetaxel active treatment start to end" and "time from post-docetaxel active treatment discontinuation to death". That said, the ERG mistakenly applied the COU-AA-301 survival (from post-docetaxel active treatment start to death) to the phase "from post-docetaxel treatment end to death"). Correctly running this analysis gives a close replication of the TA259 survival outcomes and the ICER decreases to £45,815 when the data is applied correctly.	The ERG performed an exploratory analysis by increasing post-docetaxel survival to resemble survival in TA259. As mentioned in the ERG report (Page 106), this was done "by adjusting the coefficients for Time from post-docetaxel treatment discontinuation to death". This was the only prediction equation available for time-to-death in the post-docetaxel phase. Therefore, in this exploratory analysis, survival was modified by adjusting the coefficient for "Time from post-docetaxel treatment discontinuation to death" only. Indeed, the time to treatment discontinuation while on third line active treatment (i.e. "time from post-docetaxel active treatment start to end") was not adjusted as the aim of this exploratory analysis was to model survival consistently with TA259 (not time-to-treatment discontinuation; TTD). Moreover, the incremental TTD while on 3rd line active treatment in the DES model provided by the manufacturer (MS Table 68: [redacted] LY – [redacted] LY = [redacted] LY) already corresponded closely to the incremental median TTD of TA259 (MS TA259, Page 59: [redacted] LY – [redacted] LY = [redacted] LY).

			This is not a factual error. Therefore, no changes have been made to the report. Moreover, the ERG could not validate (the methods to obtain) the ICER of £45,815 provided by the Manufacturer.

Issue 15 Clarification of error in docetaxel AE costs applied in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 98 If the ERG recalculates this deterministic ICER based on the economic model provided by the manufacturer, a slightly higher ICER of £46,756 was calculated (consistent among multiple re-runs). Differences between the ICER provided by the manufacturer and the recalculated ICER were due to lower on- and post-docetaxel Grade 3/4 AE costs. As a result, the incremental Grade 3/4 AE costs went from -£21 to -£2 (in both cases Grade 3/4 AE costs were higher for BSC).	Janssen respectfully request this statement is removed, due to typographical error in the model prior to submission.	A slightly higher ICER was generated by the ERG on multiple re-runs because docetaxel AE rates were wrongly overwritten prior to the model being supplied to the ERG on the “clinical inputs” worksheet (H120:H133). Janssen apologizes and acknowledges this was a typographical mistake prior to submission, and has resulted in docetaxel AE costs being double counted when the ERG ran the model.	We have recalculated the deterministic ICER based on the economic model provided by the manufacturer, which resulted in a slightly higher ICER. This is not a factual error. Therefore, no changes have been made to the report.

References

Kirby M, Hirst CJ, Cabrera C. Epidemiology of castration resistant prostate cancer: a longitudinal analysis using a UK primary care database. Cancer Epidemiol 2012 Dec;36(6):e349-e353.

Ryan CJ, Morris M, Molina A. Association of radiographic progression-free survival (rPFS) adapted from Prostate Cancer Working Group 2 (PCWG2) consensus criteria (APCWG2) with overall survival (OS) in patients (pts) with metastatic castration-resistant prostate cancer (mCRPC): results from COU-AA-302. Annals of Oncology 2012;23(Supplement 9)