Janssen's Response to the Appraisal Consultation Document

Abiraterone acetate for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen

Please find below Janssen's response to the Appraisal Consultation Document (ACD).

We are pleased to have the opportunity to provide our comments in relation to the interpretation of the cost-effectiveness evidence within the ACD and these are detailed below.

Our detailed response to the ACD is split into three main sections. In section 1, we provide our comments on key areas that were highlighted by the Appraisal Committee and which are material to the assessment of clinical and cost-effectiveness. In section 2, we provide a revised set of economic analyses based on a revised patient access scheme that the company are proposing. In Section 3 we respond to the remaining questions posed as part of the consultation.

Section 1

Has all of the relevant evidence been taken into account and are the summaries for clinical and cost-effectiveness reasonable interpretations of the evidence?

We have a number of comments relating to these aspects of the appraisal. Following agreement from NICE, Janssen are submitting new evidence to address some of the Committee's key concerns outlined in the ACD. In summary, the main points we would wish to draw to the attention of the Committee, and which we have provided additional evidence, are as follows.

- The licensed indication for abiraterone acetate in metastatic castrate resistant prostate cancer (mCRPC) covers the same eligible population as cabazitaxel, which was considered to meet end of life criteria. We believe that evidence presented in this response demonstrates that all three end of life criteria are met.
- The use of abiraterone acetate in the 'One Prior Chemotherapy' population is representative of those who will receive the treatment in UK clinical practice. In this response, we are pleased to have the opportunity to demonstrate with new analyses of the clinical trial that there is a sound, biologically plausible basis for the selection of this population, which demonstrates differential clinical and cost-effectiveness outcomes compared to the whole population.
- The utilities applied in the pre-progression health states within the economic analysis have been derived from robust analyses and are aligned with those accepted in other Technology Appraisals for metastatic and advanced solid tumours. Whilst the choice of

post-progression utility comes from a separate source the model is insensitive to a wide range of post-progression utility values.

- The use of Kaplan-Meier data in this case is an appropriate approach to analysing the overall survival and progression free survival data. As no single approach appears to explain the data, the Kaplan Meier approach taken in our base case can be considered closer to what could be expected in England and Wales for this patient population, given the maturity and completeness of the data in the COU-AA-301 study.

A detailed response to each of these points is provided on the following pages.

Section 4.19: The Committee concluded that abiraterone was not licensed for a small population, and therefore considered that it does not meet the criteria for an end-of-life treatment.

To be eligible to be considered under end of life criteria, the technology under review, must fulfil all three of the stated criteria. The ACD states that '*the Committee agreed that the first criterion related to life expectancy was fulfilled*' and that *'a mean improvement of greater than 3 months in mean overall survival had been robustly demonstrated*'. However, on this final criterion, the Committee concluded that '*abiraterone was not licensed for a small population, and therefore considered that it does not meet the criteria for an end-of-life treatment'.* In the following sections, we detail the reasons why we believe that abiraterone acetate does meet the criteria in terms of being licensed for a small population.

Janssen believes that it is inconsistent for NICE to conclude that cabazitaxel meets end of life criteria but abiraterone acetate does not, when both products are licensed for use in the same patient population.

Both abiraterone acetate and cabazitaxel are licensed for use in patients with mCRPC who have previously received docetaxel and so, within the terms of the license, it would be consistent to conclude that the incidence and prevalence of patients on which eligibility is based is identical. The information presented below provides evidence to support that this patient population is indeed small enough to meet end of life criteria.

During the meeting, it appeared that the views of the commissioner from Bedfordshire and Airedale PCT had overestimated the number of eligible patients. The patient number estimates suggested by the commissioner are taken from the NICE docetaxel costing template (TA101) and reference the 2008 incidence of prostate cancer across the whole of the UK (n=37,051) rather than just for England and Wales (n=33,373). With regards to the commissioner's estimation of the incidence of mCRPC, we would like to point out that whilst the incidence of prostate cancer has increased between 2006 and 2008, the incidence of mCRPC is unlikely to have increased at a similar rate. Even though there has

been a rise in prostate cancer incidence in recent years, as reported by Cancer Research UK, this has not been reflected in mortality rates.[1] Much of this increase can be attributed to the incidental discovery of earlier stage prostate cancers following transurethral resection of the prostate and, more recently, the use of PSA testing.[1] Mortality estimates in prostate cancer are often used as a proxy for mCRPC due to the short survival of patients at this stage of disease and the most recent UK mortality data shows only a slight increase from 8,506 in 2006 to 8,842 in 2009[1]. As mortality figures have only increased slightly since the release of the docetaxel costing template, it can be assumed that the incidence of mCRPC has similarly remained unchanged and, therefore, it is our belief that the commissioner has overestimated the number of patients eligible for abiraterone acetate.

Secondly, we believe that the mCRPC patient numbers who have received docetaxel presented in our original submission are aligned with those presented in the cabazitaxel submission. Within the manufacturer submission for cabazitaxel,[2] the manufacturer proposes that patient population eligible for cabazitaxel is small (estimated <2000 patients in England and Wales). Sanofi-Aventis estimate that only 3,523 patients with mCRPC receive docetaxel, which is slightly lower than our original estimate of 4,300 patients. The reference for these estimates is redacted in their submission, however as Sanofi-Aventis are the manufacturer of docetaxel, it can be expected that they would have access to accurate estimates regarding the number of patients receiving this treatment. As the two medicines are licensed for the same population, Janssen contends that the estimates provided by Sanofi-Aventis are generally applicable to abiraterone acetate.

Not all patients receiving docetaxel would go on to receive abiraterone acetate, as some patients die on treatment or progress quickly after finishing treatment. In a personal communication with the authors of the TAX327 docetaxel study it was estimated that approximately 4% of patients died whilst on treatment or within two months of stopping their chemotherapy (37/1006 patients). Clinicians at the Appraisal Committee meeting also stated that if any patients were deteriorating rapidly after their chemotherapy or had an ECOG status of >2, they would not be eligible for abiraterone acetate. Furthermore, some patients may not be considered eligible for treatment due to co-morbidities such as uncontrolled hypertension, clinically significant heart disease or cardiac ejection fraction measurement of < 50%.

Following the ACD, we have consulted with four oncologists who have provided their opinion on the percentage of patients who would be eligible for abiraterone acetate following treatment with docetaxel. The responses varied from 55% to 85% (individual responses are collated in Appendix 1) and, therefore, we have assumed the midpoint of 70% of patients would be eligible for treatment with abiraterone acetate in Table 1 below. Applying this estimate to the number of patients that are expected to receive docetaxel

[1] Figures provided by the National Cancer Data Repository Prostate Cancer Analysis , South West Public Health Observatory (Apr 2011)

from the cabazitaxel appraisal, we estimate that a total of 2,466 patients are currently eligible for abiraterone acetate in England and Wales (see Table 1).

Table 1. Patient flow illustrating the number of patients that are eligible to receive abiraterone acetate

	2011	2015
Estimated number of mCRPC patients in England and Wales*	10,856	11,161
Number of mCRPC patients that receive docetaxel estimated from the cabazitaxel NICE submission[2]	3,523 (32%)	4,464 (40%)
Number of men eligible for abiraterone acetate (excludes those with ECOG >2, those deteriorating quickly or with co-morbidities)**	2,466 (70%)	3,124 (70%)

* In 2006, NICE estimated that were 10,448 men with mCRPC in England and Wales; 0.0195% of the population;[3] using 2011 population estimates,[4] this equates to 10,856 men and could be expected to increase to 11,161 in 2015.
** based on oncologist opinion

Overall, based on these epidemiological figures, we believe that it has been clearly demonstrated that abiraterone acetate is licensed for a small patient population and, therefore, does meet all three criteria as an end-of-life treatment. Furthermore, the patient number estimates are consistent with those treatments whereby end-of-life criteria has also been deemed to have been met during the past two years, for example, sunitinib, lenolidomide, lapatinib, trastuzumab, pazopanib and everolimus (see Appendix 2 for further details).

Finally, it is possible that the Committee was concerned that abiraterone acetate may not be eligible for end-of-life considerations because of the potential that it will be licensed in future for earlier lines of metastatic prostate cancer and are, therefore, making a judgement that the population will be broader than covered by the population in this appraisal. Any potential future indications remain uncertain as this time, with Phase III trials on-going with planned interim analyses. Any new license is not anticipated until [redacted] at the earliest, with a NICE appraisal occurring some time after. As with other appraisals, we accept that expansion of the indication could cause a reappraisal of the terms of end-of-life, but we believe that is an issue for NICE consideration at the time of the next appraisal.

Section 4.10: The Committee concluded that whilst "The Committee noted that this subgroup did not match the population for which abiraterone is licensed (the license does not stipulate only one prior chemotherapy) but likely reflected the population in England and Wales for whom abiraterone would be considered" it was not appropriate to restrict the data considered in the base case economic analysis to the subgroup with one prior chemotherapy.

The Committee did acknowledge the generalisability of the 'One Prior Chemotherapy' population and states *"this subgroup did not match the population for which abiraterone is*

licensed… but likely reflected the population in England and Wales for whom abiraterone would be considered." However, on statistical grounds the Committee considered that as the tests for interaction were not statistically significant for the 'One Prior Chemotherapy' population (although numerically different) with respect to the relative overall survival (OS) benefit (Section 3.3) it was not appropriate to restrict to this subpopulation for the cost-effectiveness analysis.

Janssen believes that there is strong clinical and statistical justification for the use of the 'One Prior Chemotherapy' population in the economic analysis.

The Committee had concerns that there was no statistically significant difference with respect to overall survival between the 'One Prior Chemotherapy' population and those who had had two prior chemotherapies. It is important to note that tests for interaction are generally underpowered and therefore broader evidence should be taken into consideration alongside the:

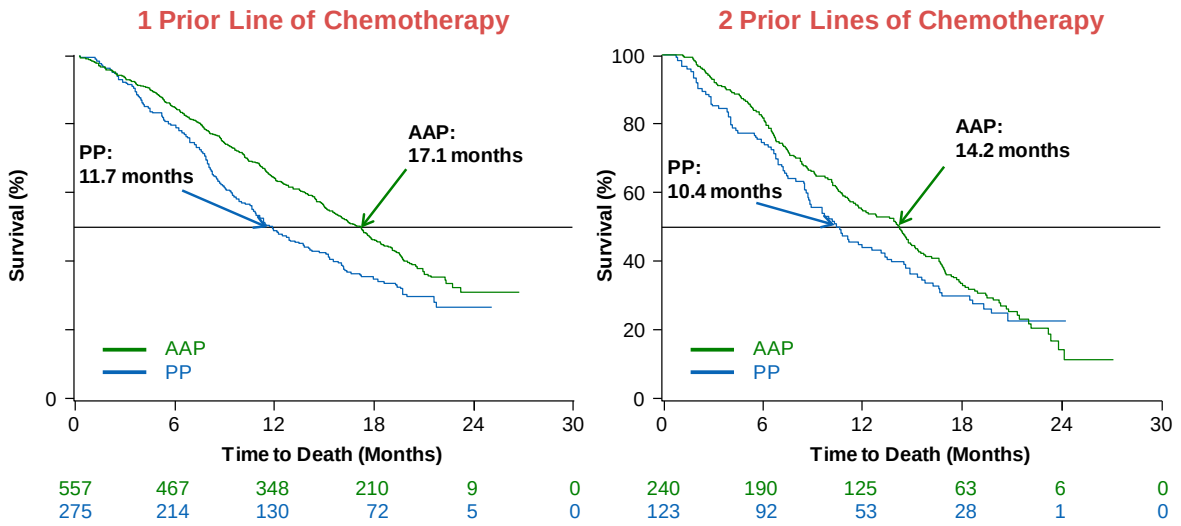
- biological plausibility of any difference in outcomes alongside pre-specification[5]
- totality of the evidence for the effect, and
- prognosis of patients and, therefore, the absolute benefit of treatment as this is also an important driver of cost-effectiveness.

In addition, the NICE Guide to the Methods of Technology Appraisal (2008) states that "it is important to consider how clinical and cost effectiveness may differ because of differing characteristics of patient populations" p.47.

Biological plausibility and pre-specification

It is important to highlight that the number of prior chemotherapies (1 vs 2) was a stratification factor in the COU-AA-301 study as it was seen as an important prognostic factor that potentially would also affect treatment outcome on the basis that prior lines of therapy implies a later stage of the disease, worse cumulative side effects, and selection of more resistant and aggressive tumour clones. Due to its importance, this stratification factor was included in the subgroup analyses that have been presented in the clinical study report, regulatory submission and the primary publication.[6] In addition, presentation of the updated analysis at ESMO in September 2011[7] included analysis by prior lines of chemotherapy, see Figure 1 below. These results show that those with two lines of prior chemotherapy experience less survival benefit than those with only one prior line.

Figure 1. Survival Benefit Observed for Subgroups With 1 or 2 Prior Lines of Chemotherapy at Study Entry.[7]



Biological plausibility would also be supported by a consistent direction of effect for this subgroup on other clinical endpoints. Following comments from the Appraisal Committee we have explored this issue further by undertaking additional analyses of key secondary endpoints in the COU-AA-301 study. These analyses show that abiraterone acetate is more effective in the 'One Prior Chemotherapy' group than in the group receiving two lines for OS, progression free survival (PFS) (time to treatment discontinuation), modified PFS, skeletal related event (SRE) progression, and prostate specific antigen (PSA) progression. PFS shows statistical significance (p=0.0393) and PSA progression approaches statistical significance for interaction terms (p=0.0613) (see Table 2 and Appendix 3).

Table 2 Hazard ratios and interaction terms for secondary endpoints from COU-AA-301

	1 Prior Chemo	2 Prior Chemo	Difference between groups
PFS (Time to Tx Discontinuation)			
Hazard Ratio	[redacted]	[redacted]	[redacted]
Interaction term			[redacted]
Modified PFS			
Hazard Ratio	[redacted]	[redacted]	[redacted]
Interaction term			[redacted]
Radiographic progression			
Hazard Ratio	[redacted]	[redacted]	[redacted]
Interaction term			[redacted]
SRE progression			
Hazard Ratio	[redacted]	[redacted]	[redacted]
Interaction term			[redacted]
PSA progression			
Hazard Ratio	[redacted]	[redacted]	[redacted]
Interaction term			[redacted]

External data

To validate the findings seen in the COU-AA-301 re-analysis, we have examined the literature to understand whether these findings are consistent with other mCRPC studies. With regards to relative effects on OS, our original systematic review identified two other studies that have demonstrated a survival benefit in this population. Both of these studies showed similar numeric differences in the OS hazard ratio regarding one prior line versus more than one prior line; the TROPIC study comparing cabazitaxel and mitoxantrone (HR=0.67 (95% CI 0.55-0.83) vs. HR=0.75 (95% CI 0.55-1.02) respectively) and the AFFIRM study comparing MDV3100 vs. best supportive care (HR=0.59 (95% CI 0.48-0.7) vs. HR=0.74 (95% CI 0.54-1.03) respectively). When combined with the evidence from the COU-AA-301 study these studies provide strong support for a reduced responsiveness to treatment amongst patients having received multiple lines of chemotherapy and, hence, support the importance of relying on evidence for a group of patients consistent with the patients expected to be treated with abiraterone acetate.

Prognosis and absolute effect size

For the COU-AA-301 study, table 8 of the updated clinical study report for COU-AA-301[8] shows that the prognosis by number of prior chemotherapies is indeed statistically significantly different, with the 'One Prior Chemotherapy' population having a 25% lower risk of death than the two prior lines population. We have, therefore, investigated the

mean survival differences between these populations further as these should reflect both differences in relative effect but also prognosis. The area under the curve was explored using a parametric fit to the data (Weibull) for convenience and as preferred by the ERG, details of this analysis can be found in Appendix 3. Overall survival is greater in the 'One Prior Chemotherapy' population than for the remaining patients (118 vs. 26 days), and as expected this numerical difference is far more striking than when considering hazards ratios alone.

These additional analyses, combined with evidence from other studies in the same patient population provides a justification for the use of the 'One Prior Chemotherapy' data in the base case economic modelling. As the Committee accepted that the 'One Prior Chemotherapy' population is more relevant in clinical practice in England and Wales and there is strong biological (validated both internally and externally), and economic rationale, we believe that this data set should be the primary focus for the economic evaluation.

Section 4.14: The Committee suggested that in UK practice patients may have lower baseline utility and shorter life expectancy than seen in the COU-AA-301 trial. The Committee also noted that the mapping algorithm resulted in pre progression utility values which were similar to or higher than utility values observed in the age-matched general population (Kind et al. 1998) and were higher than those used in the appraisal for cabazitaxel.

The utility values used in the model were mapped from a validated, widely used instrument (FACT-P) that was captured during the COU-AA-301 study. This study was judged by the ERG and the Committee to be robust and mature. The ERG stated that the mapping exercise to map FACT-P data collected in the study to EQ-5D has "advanced available procedures and identified serious errors in the method previously available in the public domain". The value, therefore can be considered a plausible estimate that is representative of patients at this stage of their disease. The Committee's question appears to be whether these values could be expected in the treated population. There are three important factors which we believe can reassure the Committee in this regard.

Firstly, we would like to address the perception that it is counterintuitive that a cancer patient could have utility values similar to, or above, those of the age matched population. This was well addressed, we believe, by one of the clinical advisors present at the Committee meeting who highlighted that because patients needed to be "fit enough" to receive first line chemotherapy, that it is in fact wholly plausible that they would have a health related quality of life similar or higher to that of the average man in the general population as they are unlikely to suffer significant co-morbidities often seen in a population of this age.

Secondly, we have provided some further context on the external literature. As we stated in our submission the pre-progression utility value generated from our mapping exercise

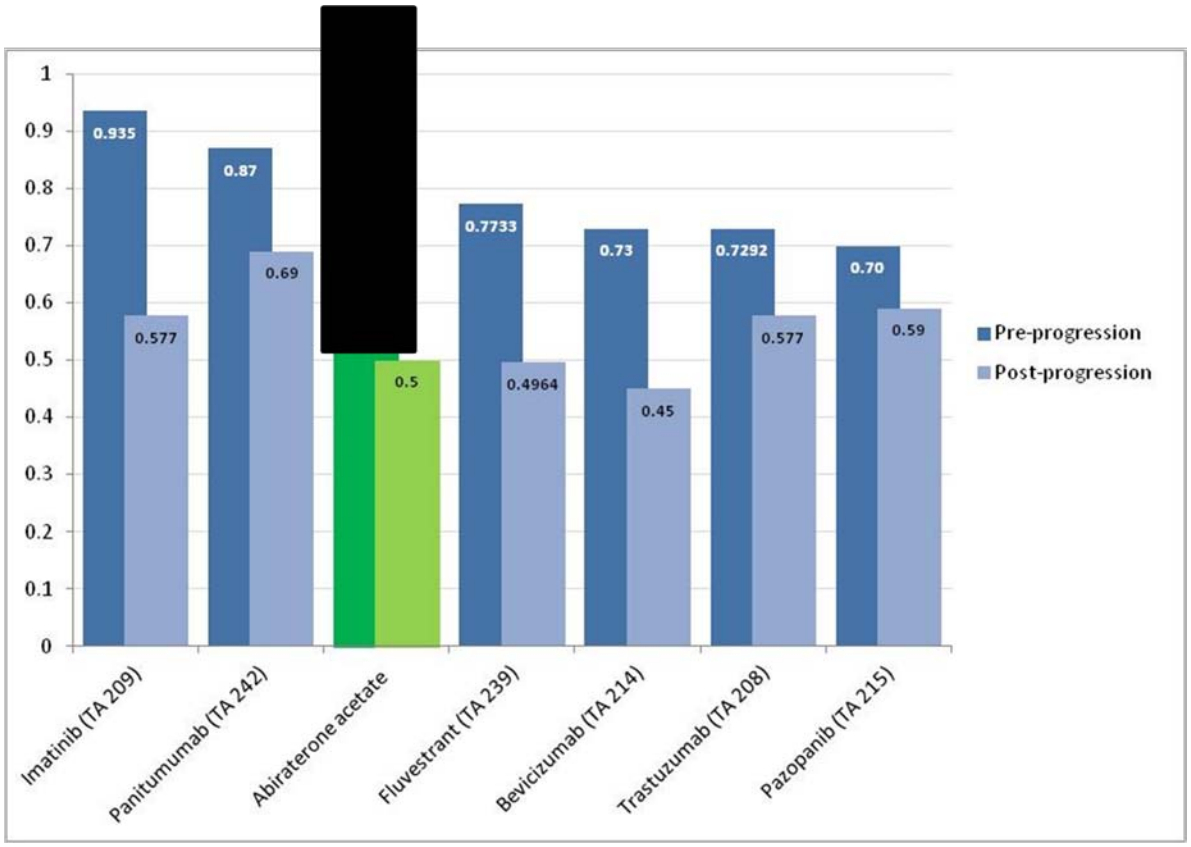
(redacted) closely matches other published data sources for this patient population (Sandblom et al., (2004) and Sullivan et al (2007)) which were not clinical trial patients. What is more, it should be noted that the data sources in the literature are often obtained from populations including high proportions of patients on chemotherapy and without regard to their progression status. In contrast, patients in the COU-AA-301 study were not taking chemotherapy (but were fit enough to have received it) and can be considered to be progression free whilst being successfully treated (pre-progression) with abiraterone acetate.

Thirdly, we would add that the demographics of patients within the COU-AA-301 study are expected to be similar to those who would receive abiraterone acetate in England and Wales who have experienced only one prior line of chemotherapy. Clinical opinion suggested that very few patients in the UK would receive more than one prior line of chemotherapy prior to receiving abiraterone acetate. We also refer the Committee to the submission for cabazitaxel which derived its utility values from an early access program in the UK and which demonstrated remarkably similar utility values. This is all the more supportive when we consider that early access programs typically recruit those patients with the highest medical need, rather than a less compromised population. Together we believe that this evidence should reassure the Committee that the values elicited directly from the COU-AA-301 study are applicable to the expected treated population in England and Wales.

Section 4.15: The Committee suggested that a smaller difference in utility between the pre-progression and post-progression health states should be assumed. The ACD states that patients in the model already have metastatic disease and further progression to the post-progression state (defined only by stopping medication) would unlikely be associated with the drop in utility modelled by the manufacturer. Furthermore the Committee criticised the fact that the pre and post utility values came from two separate sources.

The COU-AA-301 study did not collect regular quality of life information after patients had experienced disease progression. As a result of this, the only option was to use utility values for the post-progression health state from a separate data source (published literature) compared to the pre-progression utility values obtained from the COU-AA-301 study. To assess the external validity of the difference between pre- and post progression utilities, we have compared the values in this appraisal to those used in other Technology Appraisals for metastatic and advanced solid tumour cancers. The results of this comparison are shown in Figure 2 (see Appendix 5 for further details).

Figure 2: Difference in utility values between pre and post progression utility in recent metastatic and advanced solid tumour cancer appraisals by NICE.



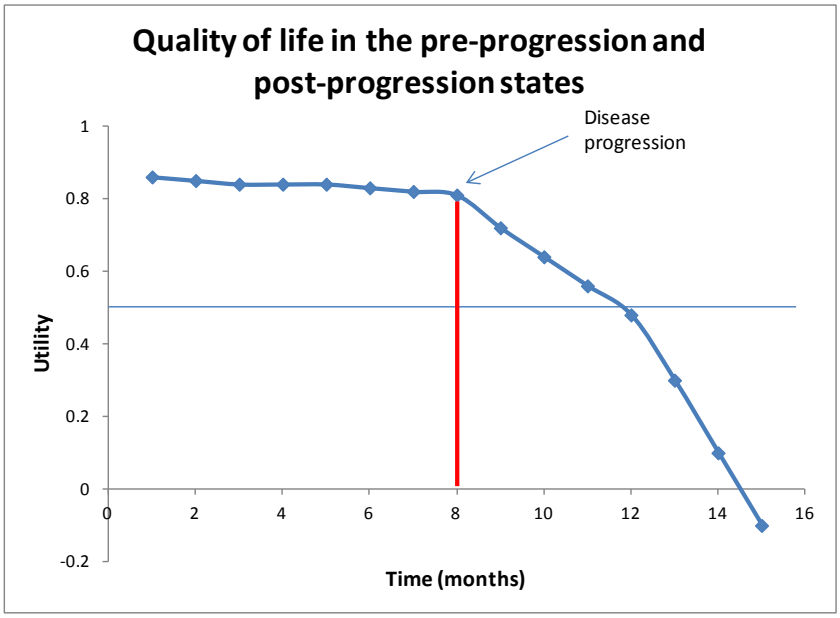
All of the appraisals detailed in Figure 2 have used post-progression utility values between 0.45 and 0.6 resulting in differences in pre and post-progression ranging between 0.11 and 0.358 when considering of the cost-effectiveness of treatments used to treat metastatic and advanced solid tumour cancers. The utility difference used for the pre- and post-progression states for abiraterone acetate is [redacted], which is in line with these other appraisals, and should be considered in the context of the acceptable tolerability profile of abiraterone acetate compared to many other cancer therapies.

The ERG suggest that one approach is to apply the mapping algorithm to those who report FACT-P at the end of treatment, which results in much higher utilities for both arms of the COU-AA-301 study (0.7 vs. 0.65 for AAP and PP respectively) instead of using the Sandblom reference (utility post progression of 0.50). However, there are a number of clear concerns with this approach.

Firstly, FACT-P data were only available for a small subsample at this point (107 and 62 in the AAP and PP arm respectively) and it is highly unlikely that those who completed the questionnaire were a random sample; it is more likely that they were a relatively 'well' group who were motivated and healthy enough to return for a post-treatment discontinuation visit. The ERG's approach of using the unadjusted estimates from these patients will have overestimated the quality of life following progression. Furthermore, this estimate is required to apply from progression until death, so again using an estimate at the point of treatment cessation will overstate the quality of life for the health state as it takes no account of the inevitable decline as a patient approaches death. Hence the approach taken by the ERG is likely to have overstated the utility following progression.

Figure 3 below illustrates the period the two utility values represent and provides the rationale for the assumption that utility value for the post-progression state is on average 0.50. After a patient's disease progresses, a steady decline in quality of life would be expected. During the last months of life a patient's quality of life would be expected to be very poor, potentially even worse than death, due to pain and debilitation associated with metastases to the bone. The utility value that needs to be incorporated into the model is not the value that is seen at the point of progression, but rather the mean utility over this period prior to death.

Figure 3. Illustrative example of quality of life in the pre-progression and post-progression health states.



It is also worth noting that the economic model is relatively insensitive to the post-progression utility value, as was demonstrated in sensitivity analyses conducted by the ERG. Therefore, although there is uncertainty around what the true value is, sensitivity

analyses showed that across a wide range of values around 0.50 value that the economic model results did not change to any great extent (p.134 of our original submission and Table 4 in this document).

In summary, the pre-progression utility values used in the model are robust as they are derived directly from the COU-AA-301 study and are aligned with several other sources in the literature within prostate cancer and across other Technology Appraisals of metastatic and solid tumour cancers. Furthermore, whilst the economic model is insensitive to the post-progression utility value, the value of 0.50 appears valid considering the quality of life that could be expected for these patients across the entire post-progression phase.

Section 4.11 and 4.13: Applying a Weibull distribution to the curves for overall survival and progression-free survival throughout the course of follow-up, and not only beyond a given time point

The Committee appear to support the ERG's view that it is more appropriate to estimate probabilities of death using survival functions throughout rather than to use the observed data from the COU-AA-301 study.

Janssen believes that there is more than one appropriate approach to the fitting of survival functions, and that additional analysis conducted fails to indicate any one 'clear cut' approach.

We note that the Guide to the Methods of Technology Appraisal (2008)[9] does not specify the need to use parametric survival functions for extrapolation in preference to using the observed data. On page 28 of the ACD, the Committee assert that 'a well-fitting parametric distribution would be more generalisable to all patients for whom abiraterone may be a potential therapy'. We understand that the Committee has to take a view on which is more appropriate, but we also believe that the choice is not clear cut and inevitably has to be a somewhat subjective judgement.

We would request that the Committee acknowledges that the functional form cannot be unequivocally chosen and would like to point out that its limited number of parameters cannot, therefore, fully describe the underlying clinical process with complete certainty. We note that the ERG did not provide any clear evidence to support their use of a parametric survival curve rather than the observed data in this case other than the DSU technical report by Latimer,[10] which suggests that parametric models for the extrapolation should be preferred 'unless data is almost entirely complete'.

In the case of the COU-AA-301 data, the survival and progression data is indeed almost complete. Therefore, a counterview is that use of a parametric model (which is unlikely to

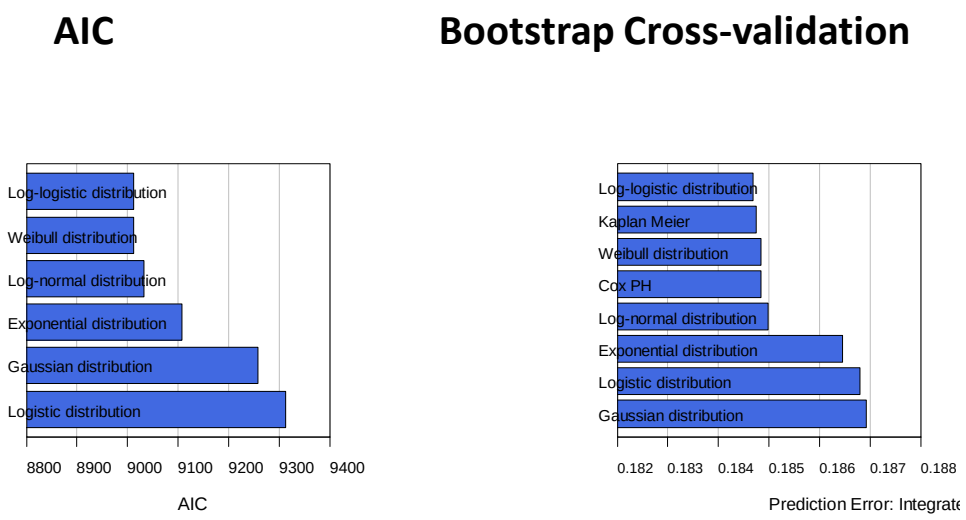
be the true specification, for this population) is unnecessary, and may introduce bias into the ensuing estimation of the QALY gain if not properly specified. Indeed in the Latimer report,[10] 38% of Technology Appraisals used a restricted means approach and almost all of these studies had survival data that were less mature than in the present submission. Fitting a parametric function to the data introduces challenges in itself. In particular, the parametric approach is overtly focused on 'internal measures of goodness of fit' but against these criteria any parametric function such as the Weibull will have a worse fit than the approach taken, of using the Kaplan-Meier curves directly.

The real value in assuming a parametric model is to allow extrapolation for those patients who remain alive at the end of the study and allow data driven assumptions about the evolution of risk of death in that period. In the economic model, Janssen has explored two approaches to this firstly, assuming the risk was constant over time for this period and secondly applying a Weibull curve which assumed a continuing increase in the risk of death.

Janssen appreciates that the ERG have assiduously investigated this issue and we appreciate their response to our comments on the ERG report. To aide their evaluation further, since the Appraisal Committee meeting, we have performed several additional analyses to further explore the appropriateness of fitting parametric functions to the data from the COU-AA-301 study. These analyses are summarised below.

The purpose of the first analysis was to evaluate the fit of different survival curves to the COU-AA-301 OS and PFS data for abiraterone acetate versus placebo and test the robustness of our assessment of goodness of fit using alternative and arguably better methods. Two methods were used to evaluate the models. The first method was Akaike Information Criterion (AIC (Akaike, 1974)) and the second was based on bootstrap cross validation techniques following the method of Gerd and Schumacher (2006 and 2007), which can be considered to be superior to the AIC method. Further details can be found in Appendix 5. The outputs are presented graphically in Figure 4 below:

Figure 4.



The aim of the second analysis was to evaluate external evidence which might inform the appropriate functional form for the survival model and whether there is consistency in the distribution describing survival over time across all the trials in mCRPC populations who have previously received docetaxel (as described in the network diagram of our original submission, Figure 13 p66, plus a recently presented study of the investigational product MDV3100). This was carried by comparing the fit of parametric functions to the OS Kaplan-Meier curves reported in the trials (n=8). Noting that six other studies did not report survival curves that could be analysed. Unsurprisingly, the survival distributions that best fit the data (according to residual deviance criteria) varied by study. The Weibull, log-normal and log-logistic distributions were most frequently identified as the best fit to the data. When all the data were considered simultaneously, the log-logistic model showed the best fit (residual deviance 272.9), however the Weibull and Log-normal also were similar (277.7 and 295.2 respectively), see Appendix 6. Given the similarities in model fit, and that the different distributions can result in very different extrapolations, choosing a parametric function on fit alone could dramatically impact the ICER. Given the results of this analysis, it appears that no one function appears to fit the trials in this patient population.

The ERG accepts that the true data for this patient population is likely to lie somewhere between the Kaplan-Meier data and the Weibull parametric function and that there is only 'moderate' evidence to support the notion that the risk of death continues to increase after the end of the study. Based on these considerations we would argue that:

- Given that the observed curves from the our study and external data do not unequivocally support a single parametric function and indeed can support models with

wildly different underlying forms, the study data itself represents the best estimate of survival for the within trial period.

- As the data equally support assumptions of an increasing or a decreasing risk of death in the extrapolation period our assumption of a constant risk of death represents a more appropriate base case than the Weibull model which is a (worst case) sensitivity analysis.

We would ask the Committee to consider that as no single approach appears to explain the data, that the Kaplan Meier approach taken in our base case be considered closer what could be expected in England and Wales for this patient population, given the maturity and completeness of the data in the COU-AA-301 study.

Section 2

Revised PAS and updated economic analysis:

In addition to providing new evidence to support the base case assumptions used in the model (relating to the use of the 'One Prior Chemotherapy' population, the approach used to extrapolate the Kaplan Meier data and the utility values) Janssen has amended the PAS to provide the Committee with greater certainty on the cost-effectiveness of abiraterone acetate.

We have revised the PAS and under this revised scheme, [redacted] ([redacted]% discount from list price). Following the ACD, we have made some amendments to the model base case to reflect the changes suggested by the ERG. The revised base case incorporates changes to administration costs per oncology consultation (£101), mitoxantrone acquisition cost (£187 per dose) and administration costs (£212 per dose) and changes to the percentage of patients receiving bisphosphonates post-progression (37%). The base case analysis has been re-run and has resulted in a reduction in the ICER from £52,688 (with the original PAS) to £46,800 (with the revised PAS) when AAP is compared to PP, see Table 2.

Table 3: Base case results ('One Prior Chemotherapy' population) for AAP vs the main comparator PP.

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
PP	[redacted]	[redacted]	[redacted]				
AAP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£46,800

One way sensitivity analysis

The deterministic sensitivity analyses consistently ranged between £42,904 and £51,110/QALY for AAP vs. PP (revised base case value £46,800). For completeness the ICERs for AAP vs. MP are also presented. The one way sensitivity analyses demonstrate that in the large majority of cases the ICER vs. PP remains under £50,000/QALY and is under £46,000/QALY for all analyses vs. MP. In general, the model is insensitive to manipulation of the time horizon, post-progression utility values and both scheduled and unscheduled medical resource use costs. In addition, if a Weibull parametric function is applied to the OS curves the ICER increases slightly to £49,911.

The model is more sensitive to the utility value for the pre-progression state, although the Committee can be satisfied that the value comes from a robust mapping exercise from a large, mature study. If the value is reduced from the base case value of [redacted] to 0.715 then the ICER vs. PP increases slightly to £51,110.

Table 4. One way sensitivity analysis results

Parameter	Baseline Value	Alternate Value	ICER vs. PP (£/QALY)	ICER vs. MP (£/QALY)
All parameters at baseline values			**£46,800**	**£41,598**
Time Horizon	10 years	4 years	£50,493	£44,738
		6 years	£47,867	£42,512
		8 years	£47,055	£41,817
Discount rate - costs	3.5%	0%	£48,985	£43,847
		6%	£45,403	£40,163
Discount rate - benefits	3.5%	0%	£43,535	£38,596
		6%	£49,111	£43,730
Overall survival approach	KM+10% cutoff+constant hazard projection	KM+10% cutoff+Weibull projection	£50,039	£44,488
		KM+5% cutoff+constant hazard projection	£47,998	£42,663
		KM+5% cutoff+Weibull projection	£50,764	£45,138
		Parametric (Weibull-placebo, Weibull-AA)	£49,911	£44,376
		Lower end of the 95% CI of KM	£44,871	£39,718
		Higher end of the 95% CI of KM	£49,102	£43,643
PFS approach	KM+5% cutoff+constant hazard projection	KM+5% cutoff+Weibull projection	£47,204	£41,846
		KM+10% cutoff+constant hazard projection	£47,015	£41,730
Baseline utility of mCRPC	[redacted]	0.715 (Sullivan, 2007)	£51,110	£45,652
		0.85 (Krahn, 2003)	£42,904	£37,967
Utility increment during abiraterone acetate treatment per cycle	[redacted]	0.036 (-20%)	£48,130	£42,174
		0.054 (+20%)	£45,788	£41,148
Utility of mCRPC post	0.5	0.4 (-20%)	£45,533	£40,414

progression		0.46 (Sandblom 2004)	£46,285	£41,116
		0.6 (+20%)	£48,139	£42,853
		0.70	£49,557	£44,187
Utility Grade 3/4 AEs	0.072	0.127	-	£41,033
Scheduled follow-up costs		-50%	£45,318	£40,576
		50%	£48,282	£42,620
Unscheduled, event-related MRU cost		-50%	£47,435	£42,266
		50%	£46,165	£40,930
Terminal care costs	£3640	£0	£46,909	£41,712
		(+20%)	£46,778	£41,575

Probabilistic sensitivity analysis (PSA)

The PSA is presented for the base case in Figure 5 and Figure 6 below.

Figure 5. PSA cost effectiveness scatter plot ('One Prior Chemotherapy' population)

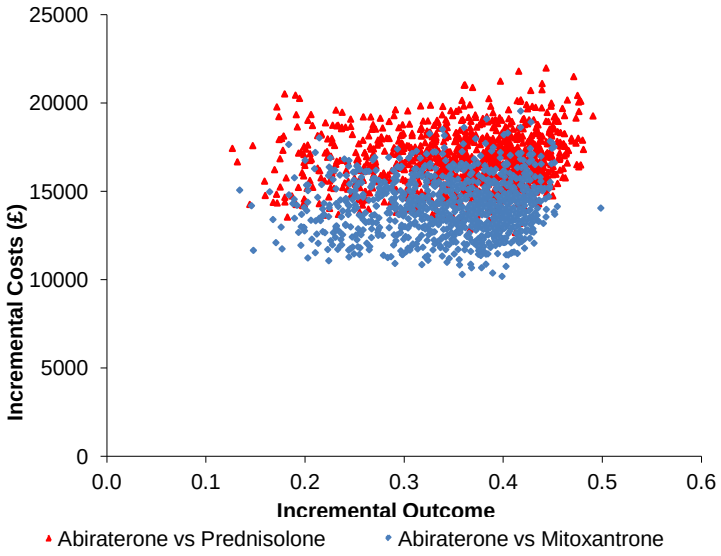
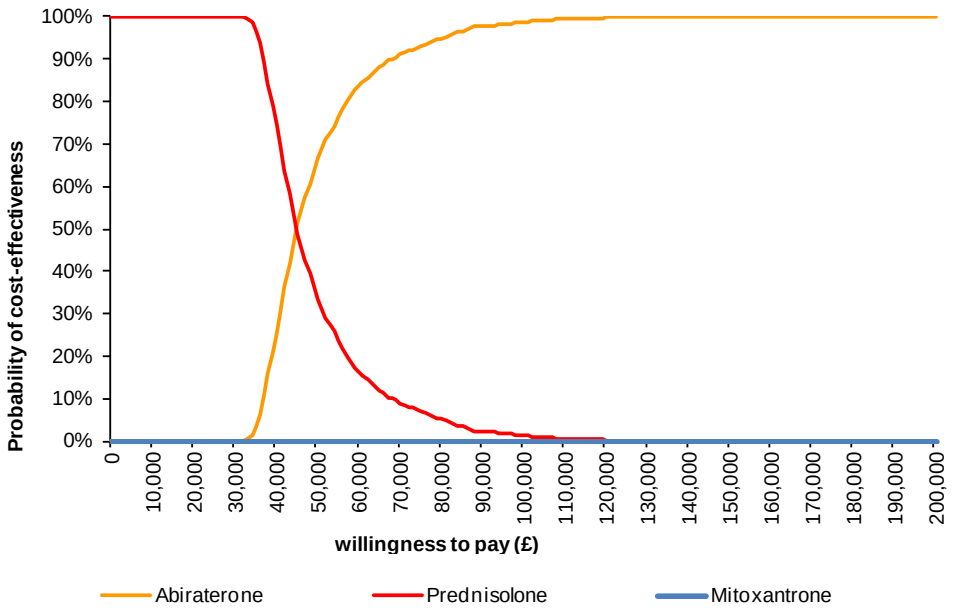


Figure 6. Cost-effectiveness acceptability curve ('One Prior Chemotherapy' population)



The probabilistic sensitivity analysis for the base case is further described in Table 4. At a QALY threshold of £50,000 the probability that AAP is the most cost-effective treatment option is 67% and at a QALY threshold of £55,000 this increases to 76%.

Table 5. Summary of PSA results

	PP	AAP
£45,000	49%	51%
£50,000	33%	67%
£55,000	24%	76%
£60,000	16%	84%

The impact of using the ITT population is explored in more detail in Scenario 1 below. We have also provided the Committee with a weighted comparator scenario (assuming 80% of patients receive PP and 20% MP), Scenario 2, to demonstrate the combined impact of introducing abiraterone acetate as a replacement for both the treatments most commonly used at this stage of the disease.

Scenario 1: ITT population

In section 1, we provided additional justification for the 'One Prior Chemotherapy' group. With the revised PAS, if the ITT population is used, the ICER is £52,851 for AAP compared to PP.

Table 6. Economic results based on the ITT population

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
PP	[redacted]	[redacted]	[redacted]				
AAP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,851

Scenario 2: Weighted average of both comparators

For the 'One Prior Chemotherapy' population we also explored the ICER comparing AAP with a basket of comparators whereby 80% of patients received PP and 20% received MP (as estimated in the clinical consensus meeting[11]) The ICER compared to the comparator basket was £45,802.

Table 7. Economic results based on a basket weighted average for PP and MP.

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER
PP/MP basket	[redacted]	[redacted]			
AAP	[redacted]	[redacted]	[redacted]	[redacted]	£45,802

The Committee had concerns about three key issues in the ACD relating to the appropriateness of the One Prior Chemotherapy population, utility values used in the modelling and the approach to modelling of the survival data. In this document we have presented further analyses and evidence to support our base case position regarding these three issues.

Janssen maintains that the 'One Prior Chemotherapy' population is more representative of the population who will receive abiraterone acetate in England and Wales and that we have provided evidence that there is a strong biological rationale for the selection of this population. In this patient population the base case ICER is £46,800/QALY and one way sensitivity analysis consistently produced ICERs less than £50,000/QALY vs. PP. The ICERs were lower still when AAP was compared with MP (base case £41,598/QALY). When a scenario was considered that explored a weighted average of PP and MP, the ICER decreased from the base case of £46,800 to £45,802/QALY. These ICERs are similar to those seen for other oncology products that have been accepted by NICE under end of life criteria.

The Committee also had concerns about the utility vales used in the modelling, whilst the model is fairly insensitive to the utility values assigned to the post-progression state, around which there is some uncertainty, there is greater certainty around the value assigned to the pre-progression health state as this data was collected from a large, mature study. We have referenced utility values that have been accepted in other appraisals of metastatic solid tumours and our values seem well within what has been accepted previously, especially given the tolerability profile of abiraterone acetate.

Finally, as no single parametric approach appears to explain the data, it appears to be an equally valid approach to use the Kaplan-Meier data from COU-AA-301 especially given the maturity and completeness of this dataset. In sensitivity analysis, application of a Weibull parametric function to the OS curves slightly increased the ICER to £49,911/QALY vs. PP.

In summary, the revised PAS has reduced the ICER in the base case to £46,800 and sensitivity analyses around some of the key areas of concern for the Committee demonstrate that the model consistently produces ICERs less than £50,000/QALY with the revised PAS. We have provided strong evidence that abiraterone acetate is indicated so treat a small patient population in England and Wales and therefore should be considered to meet all three end of life criteria.

Section 3: Response to other questions

Are the provisional recommendations sound and a suitable basis for guidance to the NHS?

Janssen propose that the Committee consider the new evidence submitted within this response, which provides a more robust, detailed, evidence-based grounding for the Committee's decision making.

Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of gender, race, disability, age, sexual orientation, religion or belief?

No

Are there any equality-related issues that need special consideration and are not covered in the appraisal consultation document?

No

Appendix 1

Thinking of your patients that have previously had docetaxel, can you estimate the number of patients that would <u>not</u> be eligible for treatment with abiraterone for each of the following reasons:

Clinician	Proportion of patients who die whilst on docetaxel or within two months of stopping treatment	Proportion of patients who are deteriorating too rapidly and would not be suitable for abiraterone	Proportion of patients with an ECOG >2 following docetaxel	Proportion of patients with uncontrolled hypertension, clinically significant heart disease of a cardiac ejection fraction <50%	Total ineligible
[redacted]	<5%	5%	5-10%	<5%	15-30%
[redacted]	5-10%	10%	10%	5-10%	30-40%
[redacted]	20%	15% not fit enough		<5%	30-40%
[redacted]	<5%	<5%	35%	<5%	35-45%

Appendix 2 Other appraisals where end-of-life criteria has been met

TA169, Sunitinib (Renal cell carcinoma) March 2009	Sunitinib is recommended as a first-line treatment option for people with advanced and/or metastatic renal cell carcinoma who are suitable for immunotherapy and have an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1. **Estimated patient numbers 2438**.
TA171, Lenolidomide (Multiple myloma) June 2009	Lenalidomide in combination with dexamethasone is recommended, within its licensed indication, as an option for the treatment of multiple myeloma only in people who have received two or more prior therapies, with the following condition. **Estimated patient numbers 2100.**
Lapatinib (Advanced/ metastatic breast cancer) FAD May 2010, appraisal suspended	Lapatinib, in combination with capecitabine, is not recommended for the treatment of women with HER2-expressing, advanced or metastatic breast cancer that has progressed following treatment with anthracyclines, taxanes, and trastuzumab in the metastatic setting, except in the context of clinical trials. **Estimated patient numbers 2000.**
TA208, Trastuzumab (HER2-positive metastatic gastric cancer) Nov 2010	Trastuzumab, in combination with cisplatin and capecitabine or 5-fluorouracil, is recommended as an option for the treatment of people with human epidermal growth factor receptor 2 (HER2)-positive metastatic adenocarcinoma of the stomach or gastro-oesophageal junction who: - have not received prior treatment for their metastatic disease and - have tumours expressing high levels of HER2 as defined by a positive immunohistochemistry score of 3 (IHC3 positive). **Estimated patient numbers 7000.**
TA215, Pazopanib (advanced renal cell carcinoma) February 2011	Pazopanib is recommended as a first-line treatment option for people with advanced renal cell carcinoma: - who have not received prior cytokine therapy and have an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 and - if the manufacturer provides pazopanib with a 12.5% discount on the list price, and provides a possible future rebate linked to the outcome of the head-to-head COMPARZ trial, as agreed under the terms of the patient access scheme and to be confirmed when the COMPARZ trial data are made available **Estimated patient numbers <4000.**
TA219, Everolimus (advanced renal cell carcinoma) April 2011	Everolimus is not recommended for the second-line treatment of advanced renal cell carcinoma **Estimated patient numbers <4000.**

Appendix 3 Additional analyses on the one prior chemotherapy

Table 8. Comparison of Treatment Effects By Endpoints (Time Unit=Day)

	1 Prior Chemo			2+ Prior Chemo			**Difference B/W Groups**
	AA (N=557)	Pred (N=275)	Difference	AA (N=240)	Pred (N=123)	Difference	
Overall Survival							
Hazard Ratio							
Interaction term							
Mean from Weibull							
PFS (Time to Tx Discontinuation)							
Hazard Ratio							
Interaction term							
Median							
Modified PFS							
Hazard Ratio							
Interaction term							
Median							
Radiographic progression							
Hazard Ratio							
Interaction term							
Median							
SRE progression							
Hazard Ratio							
Interaction term							
Median							
PSA progression							
Hazard Ratio							
Interaction term							
Median							

*The cutoff date for the mean survivals 671 days.

Area under the curve (mean) for overall survival was also explored. With KM data, the incremental treatment effect measured by mean OS between the One Prior Chemotherapy group and two prior chemotherapies group is 17 days (62 vs. 45), which is about 38% increase. Because a large number of patients are censored at the end of the trial period, mean OS with KM data is significantly underestimated. Using a Weibull fit, the projected incremental treatment effect is 92 days (118 vs. 26). However, the Weibull estimates project that prednisolone patients with two prior chemotherapies live slightly longer than abiraterone patients after trial period, which may not be true. Nevertheless, the area under

the curve analyses show that abiraterone is more effective in the One Prior Chemotherapy patients, and this incremental effectiveness may be clinically meaningful/significant.

Table 9. Weibull Estimates

	1 Prior Chemo		2 Prior Chemo	
	Estimate	Std. Error	Estimate	Std. Error
Abiraterone				
Intercept	6.4901	0.0411	6.2421	0.0478
Scale	0.7155	0.0355	0.6262	0.0408
Weibull shape	1.3977	0.0693	1.5971	0.104
Prednisone				
Intercept	6.2602	0.0559	6.1458	0.0851
Scale	0.7551	0.0481	0.8018	0.073
Weibull shape	1.3243	0.0843	1.2472	0.1135

Appendix 4. Utility values in other appraisals

Utility values estimates in recent metastatic cancer appraisals by NICE. Values used in the ERG review are listed, where the assessment group used manufacturer submission values this is highlighted as (MS).

	Pre-progression value	Post-progression value	Difference	Comment
Fluvestrant for the treatment of locally advanced or metastatic breast cancer (TA 239)	0.7733	0.4964	0.2769	The Committee agreed that, although the utility values were not generated in line with the NICE reference case, they probably represent the best published estimates available.
Panitumumab for metstatic colorectal cancer (TA 242)	0.87	0.69	0.18	These are utilities from the ERG evaluation. The committee considered that the utilities used in the manufacturer submission were highly uncertain.
Pazopanib for first-line treatment of patients with advanced renal cell carcinoma (TA 215)	0.70 (MS)	0.59 (MS)	0.11	The Committee agreed that the difference in utility values between the health states was reasonable and therefore accepted the utility values modelled by the manufacturer.
Bevicizumab in combination with taxanes for the treatment of HER2-negative metastatic breast cancer (TA 214)	0.73 (MS)	0.45 (MS)	0.28	The committee did not challenge the utility values used in the manufacturer submission
Imatinib for the treatment of unresectable and/or metastatic gastrointestinal stromal tumours (TA 209)	0.935	0.577	0.358	The Committee also noted that this utility value was higher than the value used in NICE technology appraisal guidance 179 on sunitinib for the treatment of GISTs after disease progression on imatinib treatment. Although the Assessment Group carried out some sensitivity analyses that varied the utility value, the Committee was not convinced that the most plausible value had been used and considered that this added further uncertainty to the model.
Trastuzumab for the for the treatment of HER2 positive metastatic adenocarcinoma of the stomach or gastro-oesophageal junction (mGC) (TA 208)	0.7292 (MS)	0.577 (MS)	0.1522	The committee did not challenge the utility values used in the manufacturer submission

Appendix 5. Evaluation of parametric fits within the COU-AA-301 using different methodologies.

Statistical Methods

Parametric survival models were fitted to the data for overall survival and progression free survival with treatment as a covariate. The distributions tested were

- Weibull
- Exponential
- Log-normal
- Log-logistic
- Gaussian
- Logistic

Models were fitted using the `rms` package (Harrell, 2011) in R (R Development Core Team (2011)).

Two methods were used to evaluate the models. The first method was AIC (Akaike Information Criterion (Akaike, 1974)) and the second was based on bootstrap cross validation techniques following the method of Gerd and Schumacher (2006 and 2007). The AIC score is an estimate of the error in the model penalised by the number of parameters. The penalisation is required to help avoid selecting models that over-fit the data. In contrast bootstrap cross-validation is an exact method of estimating the error in the model. The prediction error is determined for a series of time points during the study period and the weighted area under this curve is used to derive the (IBS) Integrated Brier Score (also known as cumulative prediction error or cumulative rank probability score). The lower this value the better the fit of the model or the more reliable it is for making future predictions. Any model can be evaluated using this technique which means that Cox Proportional Hazard model and Kaplan Meier models can be used for reference. This method of evaluation is considered superior to using Akaike Information Criterion (AIC) or *k*-fold cross-validation (Harrell, 2001). A total of 500 bootstrap samples were used to evaluate the models from the start time up to the maximum time available from the data. The package `pec` (Gerds, 2009) in R was used to conduct this analysis. Two measurements of error are provided. One is the bootstrap cross validation error and the second is the apparent error. The apparent error is the error of fitting the model to the same data from which it came and the bootstrap cross validation error is the error when fitting the model to a new data set. There are three different ways of using these estimates of error.

- Use only the bootstrap cross validation error
- Use a linear combination of the apparent error and bootstrap cross validation error using the constant weight 0.632.

- Use a linear combination of the apparent error and bootstrap cross validation error using weights dependent on how the models perform in permuted data. (Efron's and Tibshirani's 632+ estimate (Efron and Tibshirani, 1997))

The argument to combine apparent error with the bootstrap cross-validation error is that the bootstrap cross-validation error over-estimates the error since it is based on smaller samples than the original data set. For this report, only the bootstrap cross validation error was used since we are only interested in comparing models rather than trying to get an estimate of the exact prediction error. This is the recommended method of Pers *et al.* (2009) who argue that the other two methods still require more validation work on them before being used more widely.

The results showed that the parametric fits generally fitted one arm of the study well but fitted the other arm poorly. The three best fitting models were refitted using a stratified option which allows the scale parameter to differ between study arms. AIC values were calculated for these models. The results of all the models tested are presented. However, the bootstrap cross-validation method is currently not available for stratified parametric models.

Sample Size and Status of Patients at End of Study Period

The status of patients in terms of survival is shown in Table 1 and in terms of progression free survival is shown in Table 2. Length of study was approximately 2 years.

Table 1. Status of patients at end of study period (survival)

Treatment	Alive	Dead	Total	% died
Abiraterone	296	501	797	63%
Placebo	124	274	398	69%

Table 2. Status of patients at end of study period (progression)

Treatment	Not progressed	Progressed	Total	% progressed
Abiraterone	54	743	797	93%
Placebo	17	381	398	96%

Results for Overall Survival

The shape of the survival curves is shown in Figure 1.

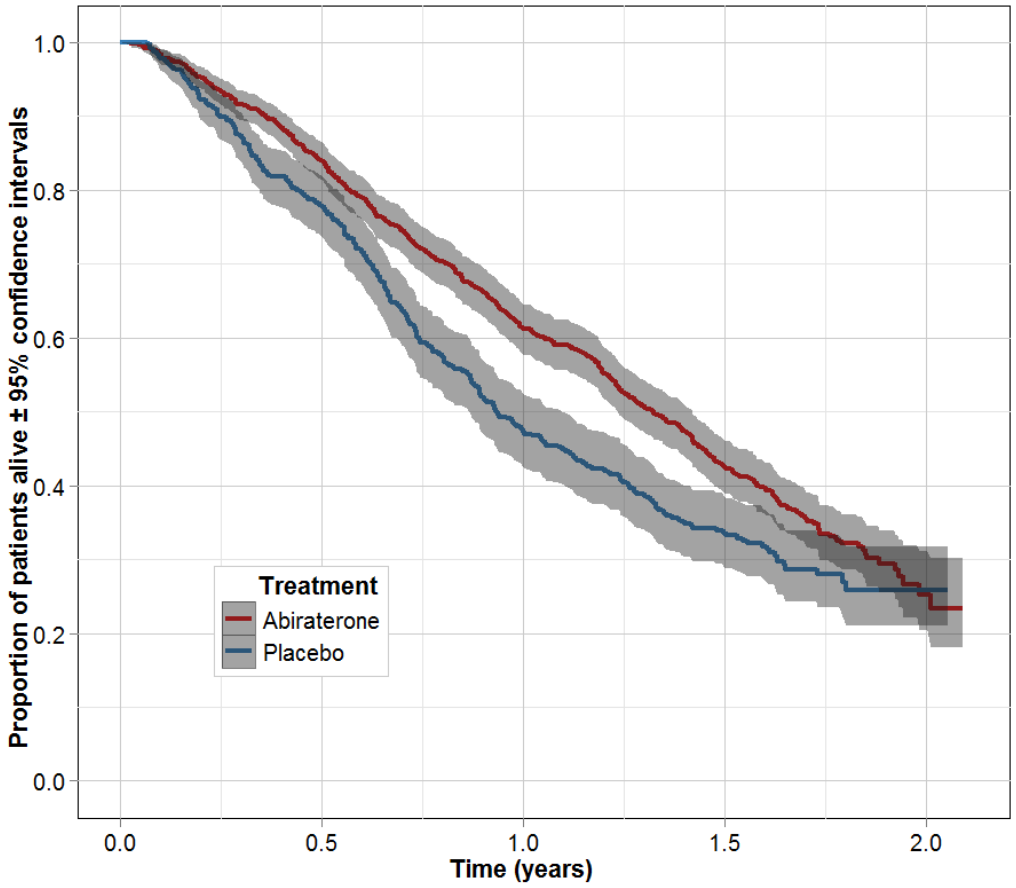


Figure 1 Kaplan Meier curves for Abiraterone and Placebo (overall survival).

Sample size at end of Kaplan Meier chart

Placebo: n = 25
Abiraterone: n = 14

Check Proportional Hazard Assumption for Overall Survival

The original report used an extrapolation based on a proportional hazard assumption. The analysis below checks this assumption. A log-log plot is shown in Figure 2. If the proportional hazard assumption holds then the chart should show 2 parallel lines. However, this is not the case.

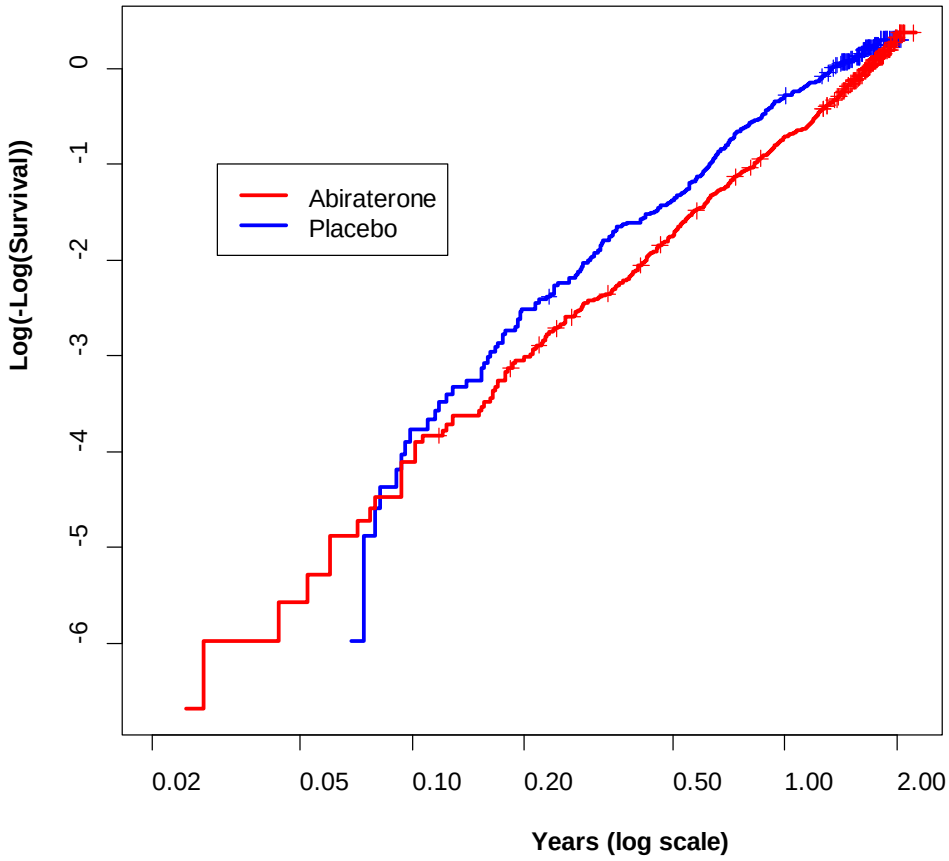


Figure 2. Log-log survival curves for Abiraterone and Placebo.

Test proportional Hazard Assumption for Overall Survival

A Cox Proportional hazard model was fitted to the data the proportional hazard assumption tested. The results are shown below, which indicate that the data deviates significantly from the proportional hazard assumption.

	Rho	Chi-sq	P
Treatment	-0.102	8.07	0.0045

AIC for Overall Survival

The AIC scores are shown in Figure 3.

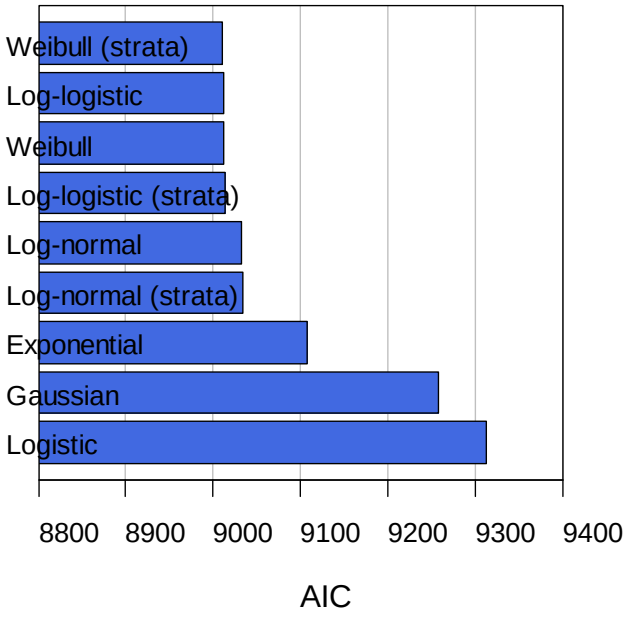


Figure 3. AIC for models fitted to the overall survival data. The lower the score the better the fit.

Bootstrap Cross-validation for Overall Survival

The IBS are shown in Figure 4.

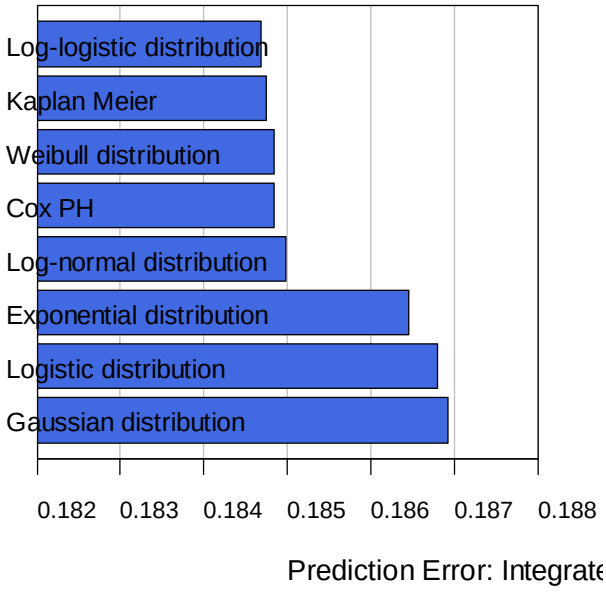
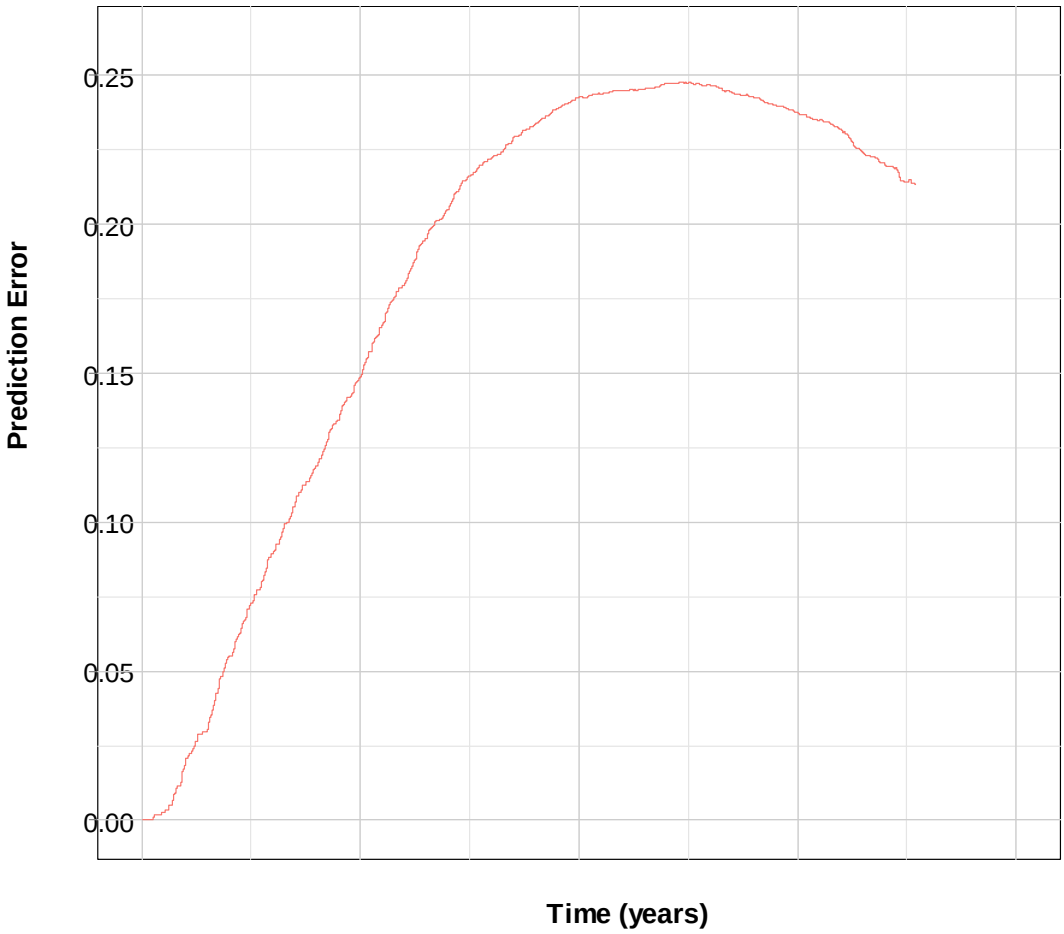


Figure 4. IBS for models fitted to the overall survival data. The lower the score the better the fit.


Prediction Error
0.25
0.20
0.15
0.10
0.05
0.00
Time (years)

Appendix 6. Evaluation of parametric fits of other RCTs in the same patient population as indicated for abiraterone

Mapi JA12230A OS data fit v2.ppt

References

1. Cancer Research UK. Prostate cancer - UK incidence statistics. 2-6-2009.

2. Sanofi-Aventis. Cabazitaxel for the second-line treatment of metastatic hormone refractory prostate cancer. 2011

3. National Institute of Health and Clinical Excellence. Costing template for NICE technology appraisal 101-Docetaxel for the treatment of hormone-refractory metastatic prostate cancer. 2006

4. Office of National Statistics. Vital Statistics: population and Health Reference Tables (Autumn 2011). 2011

5. Sculpher M. Subgroups and Heterogeneity in Cost-Effectiveness Analysis. Pharmacoeconomics 2008;26(9):799-806.

6. de Bono JS, Logothetis CJ, Molina A et al. Abiraterone and increased survival in metastatic prostate cancer. N Engl J Med 2011;364(21):1995-2005.

7. Fizazi et al., Final Overall Survival (OS) Analysis of COU-AA-301, a Phase 3 Study of Abiraterone Acetate Plus Prednisone in Patients with Metastatic Castration-Resistant Prostate Cancer (mCRPC) Pretreated with Docetaxel. ECCO-ESMO Sept 2011; 2011.

8. Johnson & Johnson Pharmaceutical Research & Development LLC. Statisical Report of Updated Data from Study COU-AA-301. 2011

9. National Institute of Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008

10. Latimer N. NICE DSU Technical Support Document 14: Surivval analysis for economic evaluations alongside clinical trials- extrapolation with patient-level data. 2011

11. GfK Healthcare. Healthcare resource utilisation and the treatment-related care-pathways within the current and future setting of metastatic castration resistant prostate cancer (mCRPC). 2011