Abiraterone for treating newly diagnosed metastatic hormone-naïve prostate cancer

Produced by Aberdeen HTA Group

Authors Ewen Cummins[1]
Lorna Aucott[2]
Clare Robertson[2]
Cynthia Fraser[2]
Rhona Johnston[1]
Thomas Lam[3]
Miriam Brazzelli[2]

1 McMDC Ltd. Health Economics, Glasgow, UK
2 Health Services Research Unit, University of Aberdeen
3 NHS Grampian, Aberdeen Royal Infirmary, Aberdeen, UK

Correspondence to Miriam Brazzelli
Health Services Research Unit, University of Aberdeen
2nd Floor, Health Sciences Building, Foresterhill
Aberdeen, AB25 2ZD
m.brazzelli@abdn.ac.uk

Date completed 10 April 2018

Version 1

Source of funding: This report was commissioned by the NIHR HTA Programme as project number **16/108/08**.

Declared competing interests of the authors

No competing interests to disclose.

Acknowledgements

The authors are grateful to Lara Kemp for her administrative support.

Rider on responsibility for report

The view expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Cummins E, Aucott L, Robertson C, Fraser C, Johnston R, Lam T, Brazzelli M. Abiraterone for treating newly diagnosed metastatic hormone-naïve prostate cancer. Aberdeen HTA Group, 2018.

Contribution of authors

Ewen Cummins acted as health economist for this appraisal: critiqued and reviewed the cost-effectiveness evidence, checked and re-analysed the economic model, and carried out further sensitivity analyses. Rhona Johnston acted as programmer and modeller and contributed to the revision of the company model and helped with some of the model rebuild. Lorna Aucott acted as statistician: critiqued the statistical methods presented in the submission, checked the numerical results, analyses, tables, and figures related to the review of the clinical effectiveness evidence. Clare Robertson acted as systematic reviewer: critiqued the company's definition of the decision problem and the clinical effectiveness evidence. Cynthia Fraser acted as information scientist: critiqued the methods used for identifying relevant studies and checked the search strategies used in the submission. Thomas Lam acted as clinical expert: provided clinical advice and general guidance. Miriam Brazzelli acted as project lead for this appraisal: contributed to the critique and review of the clinical

effectiveness methods, checked the final report and supervised the work throughout the project.

Table of contents

List of tables

List of figures

1 Summary

Prostate cancer is the most common male cancer in the UK, with over 46,700 people diagnosed in 2014. Approximately 18% of new cases present with metastases at first diagnosis, meaning the cancer is diagnosed too late for curative treatment to be possible as it has already spread outside the prostate gland and through the body. The term metastatic hormone sensitive prostate cancer refers to people who have not received hormone therapy or who have received hormone therapy but have not yet become resistant to treatment. Those with newly diagnosed metastatic hormone sensitive prostate cancer (mHSPC) have a poorer prognosis than people who are first diagnosed with localised disease.

Androgen deprivation therapy (ADT) has been the standard of care in mHSPC, orchidectomy (surgical castration) and bicalutamide monotherapy are less common treatment options. Data from two recent clinical trials, CHAARTED and STAMPEDE, have shown that the addition of docetaxel (chemotherapy) to ADT for the treatment of newly diagnosed mHSPC was beneficial in terms of health outcomes, but associated with greater toxicity and potentially severe side effects. Several novel agents are now available, such as abiraterone acetate, and the order in which a patient may receive them is determined by clinical symptoms and manifestations, prior treatment, NICE recommendation and NHS policy.

1.1 Critique of the decision problem in the company submission

The company's submission considered abiraterone acetate (trade name Zytiga) with prednisone/prednisolone (AAP) plus androgen deprivation therapy (ADT) for the treatment of adults with newly diagnosed, high risk mHSPC.

The decision problem addressed in the company's submission was broadly consistent with the NICE final scope. The NICE final scope for this appraisal specified the population as adults with newly diagnosed high risk metastatic hormone-naïve prostate cancer (mHNPC), while the population addressed in the company submission is adults with newly diagnosed, high risk mHSPC. The company state that the marketing authorisation wording describes AAP as indicated for the treatment of

newly diagnosed high risk mHSPC in combination with ADT and that the terms mHNPC and newly diagnosed mHSPC are effectively the same because newly diagnosed patients are, by default, hormone naïve. The company did not consider orchidectomy and bicalutamide monotherapy as clinical experts advised that these are seldom used in the UK. The comparators presented in the company submission are ADT alone (including LHRH agonist therapy) and docetaxel (DOC) plus ADT. The company state that clinical experts provided validation that there is no difference in the type of ADT, thus justifying their approach. The company submission includes all the outcomes listed in the NICE scope and reports additional outcomes from the LATITUDE trial: progression free survival following subsequent therapy, time to symptomatic local progression, prostate cancer-specific survival, time to chronic opiate use, castration status.

1.2 Summary of clinical effectiveness evidence submitted by the company

The clinical effectiveness evidence submitted by the company consist of one RCT, the LATITUDE trial (1199 participants), with supporting evidence of one further RCT, the STAMPEDE trial (1917 participants). LATITUDE is a manufacturer-sponsored, multinational, randomised, double-blind, placebo-controlled Phase III trial that investigated abiraterone acetate with prednisone/prednisolone (AAP) plus ADT (597 participants) versus ADT plus placebo (602 participants). The company consider the ADT plus placebo arm equivalent to ADT alone. The company also maintain that LATITUDE is the only RCT providing data specific to the target population of people with newly diagnosed, high-risk mHSPC. The manufacturer-sponsored STAMPEDE trial represents the largest evidence base of AAP plus ADT in early prostate cancer data relevant to UK practice but include a broader patient population than LATITUDE, and does not report data separately for high risk disease/high volume patients.

The co-primary outcomes assessed in the LATITIDE trial were overall survival (OS) and radiographic progression free survival (rPFS). OS was also the primary outcome in STAMPEDE whilst failure free survival (FFS) was the intermediate primary outcome. In the LATITUDE trial, treatment with AAP plus ADT was associated with a 38% reduction in the risk of death compared with ADT alone (HR=0.62 [95%CI: 0.51–0.76]; p<0.001).7 The overall survival rate at three years was 66% in the AAP +

ADT group and 49% in the ADT alone group. There was an imbalance in the proportion of patients who received life-extending subsequent therapies (20.9% in the AAP plus ADT arm versus 40.9% in the ADT alone arm). The company claim that this could result in the standard ITT analysis of OS underestimating the true OS benefit for AAP. Therefore, additional pre-specified OS analysis using the IPCW methodology were conducted by the company to adjust for patients who switched to other therapies. This analysis showed AAP plus ADT significantly improved survival compared to ADT alone, with an improved HR=0.48 (95% CI: 0.36–0.63; p<0.0001). Results from STAMPEDE are consistent with these results. Treatment with AAP + ADT was associated with a 39% reduction in the risk of death compared to ADT alone (HR= 0.61 [95% CI: 0.49–0.75]; p<0.0001).

In LATITUDE, treatment with AAP plus ADT significantly delayed disease progression compared with ADT alone. AAP + ADT resulted in a 53% reduction in the risk of radiographic progression or death (HR=0.47 [95% CI: 0.39–0.55]; p<0.001). At three years, 47% of patients in the AAP + ADT arm remained event-free, compared to only 21% of those in the ADT alone arm. In support of this evidence, the company present data from the metastatic (M1) subgroup of STAMPEDE, in which treatment with AAP + ADT was associated with a 69% reduction in the risk of biochemical failure, progression or death compared with ADT alone (HR=0.31 [95%CI: 0.26–0.37]; p<0.0001).

The median treatment duration in the safety population of the LATITUDE trial was 24 months in the AAP + ADT arm and 14 months in the ADT alone arm. Treatment emergent adverse events (TEAEs) were reported by a higher number of people in the AAP+ADT group than for ADT alone. The most frequently reported TEAEs in the (reported in ≥20% of patients) in either the AAP + ADT or ADT alone arm were hypertension (37% versus 22%, respectively), hypokalaemia (20% versus 4%) and back pain (18% versus 20%). Commonly reported serious adverse events (SAEs) (reported by ≥1% of patients in either the AAP + ADT or ADT alone group) included pneumonia (1.8% versus 0.3%, respectively), spinal cord compression (1.7% versus 1.8%) and urinary retention (1.5% versus 1.7%). The most frequently reported adverse events (AEs) leading to treatment discontinuation (reported in ≥1% of patients in either the AAP + ADT or ADT alone group) were spinal cord compression

(0.8% versus 1.0% of patients, respectively) and bone pain (0.5% versus 1.0%, respectively). Cases of discontinuation for hypokalaemia, hypertension and cardiac disorders were rare.

The comparison of the effectiveness of AAP with DOC for the mHSPC patient group was made using indirect treatment comparisons since no head-to-head studies currently exist in this particular patient group. For the co-primary outcomes, three RCTs were subsequently included: LATITUDE, CHAARTED (790 participants) and GETUG-AFU 15 (385 participants); the latter two using post-hoc selected sub-groups of newly diagnosed patients with high volume disease. STAMPEDE, which assessed a much broader patient group, was only included in sensitivity analyses.

The results suggest non-significant effects for OS (HR 0.92 [95% Crl 0.69-1.23]) and for rPFS (HR 0.76 [95% Crl 0.53-1.10]) albeit with Bayesian pairwise probabilities of 71.8% and 92.9%, respectively. These probabilities represent a level of certainty that AAP+ADT patients may be more likely to survive or have progression free survival using AAP+ADT compared with DOC+ADT. The company presented also a number of sensitivity analyses with varied but similar results.

Results of sensitivity analyses suggest that skeletal-related events (SRE) were similar in the indirectly comparison between AAP and DOC, [redacted] but with a Bayesian probability of [redacted] but without adequate group identification.

Two RCTs, LATITUDE and GETUG-AFU 15, fed into a Bayesian ITC for safety results, but no sensitivity analyses were reported. When the AAP+ADT group (n=597) was indirectly compared to the DOC+ADT group (n=189), [redacted]

The Functional Assessment Cancer Therapy-Prostate (FACT-P) and Brief Pain Inventory (BPI) quality of life measures, looked at differences of change from baseline for both AAP+ADT and DOC+ADT treatment groups over four time points 3, 6, 9 and 12 months in LATITUDE (ITT) and CHAARTED (high volume disease - HVD). Sub-group analyses were conducted by the company whereby high risk disease (HRD) and HVD patients in LATITUDE were selected post-hoc. At 3 months, AAP+ADT had a significant positive and beneficial increase on FACT-P over DOC+ADT, with difference of change = 4.20 (95% CrL 1.18-7.19) and a 99.7% probability of AAP being better than DOC. AAP estimates improved further over time as did the DOC estimates (not to the same extent and never to the level of AAP), but differences between AAP and DOC were not significant by 6 months or even at 1 year. BPI results showed larger decreases in pain estimates for indirect comparisons between AAP and DOC, but the results were not significant. Pain in the DOC group increased with time whereas with AAP they remained steady if not further reduced. The sensitivity analyses were comparable for FACT-P and BPI.

In the absence of any head-to head studies, further indirect comparisons were conducted for a group of men with disease progression (for the mCRPC group with respect to the effectiveness of AAP with other treatments including DOC). These were not presented in the clinical effectiveness section of the submission but only in the cost-effectiveness section. The company used the COU_AA_302 study, which directly compared abiratone plus prednisolone with placebo plus with prednisolone, and other studies which compared different treatments with placebo or best standard care. In particular, the company focused on DOC (the TAX327 study comparing DOC to a different placebo, mitoxantrone), radium-223 (the ALSYMPCA study with prednisolone as placebo) and enzalutamide (the PREVAIL study with prednisolone as placebo). In general, the estimates show that AAP is comparable with other treatments.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

LATITUDE has provided the only evidence so far of AAP+ADT compared with ADT alone for the treatment of men with mHSPC. The ERG agree with LATITUDE results suggesting that AAP+ADT to be beneficial for the primary outcomes of OS and rPFS and for most of the secondary outcomes of safety and quality of life compared to

ADT. In terms of safety, AAP+ADT had a slight increased risk for hypertension and hypokalaemia. The results of LATITUDE are similar to those of the STAMPEDE trial. However, the STAMPEDE patient group was broader and while the company have conducted similar analyses on a *post hoc* subgroup profiled to be similar to the LATITUDE population, they rightly have not combined the results of these studies. Overall, the results from the LATITUDE trial provide evidence of benefits of AAP+ADT over ADT alone for the treatment of patients with mHSPC for the outcomes survival, progression and quality of life. The risk of some safety outcomes increased for AAP but the ERG agree that these may be well treated medically.

With no head-to-head trials assessing the effects and safety of abiratone versus the only other relevant comparator, DOC, identified for the patient group of interest, mHSPC, indirect treatment comparisons (ITC) were a sensible option. The company used a Bayesian network meta-analysis (NMA). The primary outcomes were based on three RCTs: LATITUDE, which compared AAP+ADT to ADT alone, and CHAARTED and GETUG-ARG 15, both of which compared DOC in conjunction with ADT to ADT alone. The NMA results showed no evidence of a difference in OS and rPFS between AAP+ADT and DOC+ADT, despite the many sub-group analyses using many combinations of patient groups in an attempt to mirror the LATITUDE population. The results did not vary drastically but it is not clear which might be the most reliable.

For the relapsing/progression patients, the mCRPC group, the ITC used were Bucher pairwise estimates comparing other treatments with AAP. This approach requires many independent steps and so, intuitively, seems less robust compared to the NMA above, but the ERG agree it was probably the only course of action to accommodate the lack of studies and comparison arms. Each study compared a treatment with a ‘placebo’ although not always the same one. The conclusion that AAP is comparable to other treatments with regard to OS and rPFS is probably reasonable.

The ITC analyses for both the mHSPC and mCRPC patient groups, have basic assumption violations of contextual heterogeneity which the company discussed in some detail and acknowledge the subsequent limitations. However, no checks were provided for statistical heterogeneity or consistency. All of these mean that clinically,

the ERG agree with the company's conclusions that AAP is at least as effective as other treatments for both newly diagnosed patients and those who have relapsed or progressed. However, the decision of which estimates to use for further modelling and interpretation should be taken with caution given the spectrum of possibilities available across different credible limits.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The company compares three mHSPC treatment arms in the economic model:

- AAP+ADT
- ADT
- DOC+ADT

This also requires the company to model the treatment sequences for when patients progress to mCRPC. Which treatments patients receive for their mCRPC is determined by which treatment they received for their mCRPC. Because the LATITUDE trial is not solely UK based the company applies mCRPC treatment proportions derived from expert opinion. These mCRPC treatment proportions have some effect upon patient outcomes, but mainly affect the estimated mCRPC costs.

The company outline that all other companies submitting in the area have adopted a partitioned survival analysis. The company model is a quite complex markov model with a 20 year time horizon. Discounting and perspectives are as per the NICE reference case. The model applies the LATITUDE Kaplan Meier OS and rPFS data for the first 5 months. The LATITUDE 5 months plus data is analysed using multi-state modelling (MSM) to provide transition probabilities for 5 months plus. The DOC+ADT curves are estimated by applying the company ITC hazard ratios to the rPFS and OS probabilities in the AAP+ADT arm.

It appears that the post progression survival is divided into 1st line mCRPC treatment, 2nd line mCRPC treatment and 3rd line treatment using mean duration data from the COU-AAP-302 trial of abiraterone for mCRPC. The model that uses this method of dividing the post progression survival is referred to as the MSM model in what follows.

The model also contains the facility to apply the mCRPC discontinuation and overall survival curves estimated by the discrete event simulation model that the company submitted for TA387. These provide estimates for 1st line mCRPC treatment with abiraterone and BSC. The curves for other active treatments are estimated by applying hazard ratios to the abiraterone curves. The curves that are applied in each arm are averages of these mCRPC curves,

weighted by the arm specific 1st line mCRPC treatment proportions. For the base case the company assumes that all active treatments are equally effective for mCRPC. This determines the duration of 1st line mCRPC treatment and mCRPC survival. The mCRPC survival after discontinuation from 1st line mCRPC treatment appears to be divided into 2nd line treatment and 3rd line treatment using mean duration data from the COU-AAP-302 trial of abiraterone for mCRPC. The model that uses this method estimating mCRPC treatment and survival is referred to as the MSM/TA387

The company argue that the LATITUDE OS data are not relevant to the UK due to different treatments for mCRPC and that it is important to model the effects of these. Mainly due to this, the company prefer the MSM/TA387 model to the MSM model.

The MSM/TA387 model that applies the mCRPC discontinuation and OS curves estimated by the TA387 model results in OS curves that are a poor fit to the LATITUDE OS Kaplan Meier curves. The company fit the MSM/TA387 model OS curves to the LATITUDE OS Kaplan Meier curves by applying an ad hoc hazard ratio of 2.62 to the OS curves estimated by the TA387 model. The TA387 model discontinuation curves have a similar compensating adjustment applied. This causes the MSM/TA387 model OS curves to be aligned with the LATITUDE OS Kaplan Meier curves.

Due to the 2.62 hazard rate adjustment, the MSM/TA387 model estimates very similar OS curves to those of the MSM model during the period of the LATITUDE trial. The models' OS curves only really diverge during the period of extrapolation.

The company undertake a repeated measures analysis of the LATITUDE EQ-5D data. This estimates a treatment effect increment of [redacted] for AAP+ADT over ADT. It also estimates quite large decrements for SAEs and SREs. The decrements for SAEs and SREs are not applied. Instead the company derive smaller decrements from the literature.

The LATITUDE data do not address what the quality of life should be in the DOC+ADT arm. The company commission a TTO study from MAPI values to estimate this relative to the ADT arm quality of life. The health state descriptor for those in the DOC+ADT arm who are receiving docetaxel treatment is worse than that

for the ADT arm. RCT trial FACT-P data supports this assumption. The health state descriptor for those in the DOC+ADT arm who have completed a course of docetaxel treatment and are now only receiving ADT is also worse than that for ADT arm. This is because they are more frequently depressed. When valued by 200 members of the UK public this results in quality of life decrements in the DOC+ADT arm for those who are receiving docetaxel treatment of [redacted] and for those who have completed their docetaxel treatment of [redacted].

Drug costs for mHSPC have treatment compliance percentages applied to them. The company estimate an [redacted] percentage for abiraterone based upon the areas under the LATITUDE AAP+ADT arm rPFS and TTD curves.

Other resource use is largely based upon expert opinion. The main difference between the arms is that DOC+ADT patients receiving docetaxel are assumed to require bone scans. No bone scans are required in either the AAP+ADT arm or the ADT arm. The frequency of bone scans increases in the DOC+ADT arm when patients have completed their course of docetaxel. The number of CT scans is also slightly higher for DOC+ADT patients who have completed their course of docetaxel than for AAP+ADT patients and ADT patients.

The company base case deterministic cost effectiveness estimates are £17,418 per QALY for AAP+ADT compared to ADT and £17,828 per QALY for AAP+ADT compared to DOC+ADT. The central probabilistic estimates are aligned with these.

A range of univariate sensitivity analyses are presented which vary inputs according to their 95% confidence limits, or if these are not available by ±10%. These find results to be sensitive to the clinical and utility inputs, due in part to these having 95% confidence limits. Results are not found to be sensitive to cost inputs, but this may be due to them largely not having 95% confidence limits.

The company also present a range of scenario analyses which find results to be sensitive to:

- the time to subsequent therapy being used as the definition of progression

- the MSM model being used, with this being coupled with the LATITUDE mCRPC treatment proportions
- a time horizon of only 5 years
- applying the abiraterone quality of life increment until death
- the DOC+ADT quality of life decrement for mHSPC patients post docetaxel treatment
- vial wastage
- applying the LATITUDE QoL regression coefficients instead of the subset of the base case
- The time point of the switch from Kaplan Meier data to MSM probabilities

Some of the company scenario analyses have cost effectiveness estimates higher than £20,000 per QALY. None have cost effectiveness estimates higher than £30,000 per QALY.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

It appears that the 1st line mCRPC costs and benefits estimates of both the MSM model and the MSM/TA387 model are not reliable. All cost effectiveness estimates may consequently not be reliable.

The company cost effectiveness estimates may be biased in favour of AAP+ADT because:

- It is questionable whether there is a quality of life decrement for those who have completed a course of docetaxel compared to those who have only ever received ADT. There are reasons and trial data to suppose there may be an increment.
- If there is a quality of life decrement for those who have completed a course of docetaxel the company commissioned TTO study that estimates this may be biased.
- The company's estimates of the quality of life decrements for docetaxel are only applied in the DOC+ADT arm, and not to docetaxel treatment for mCRPC in the AAP+ADT arm or the ADT arm.

- The company only partially apply the results of the LATITUDE QoL regression, which pushes up quality of life values to above those observed during the LATITUDE trial.
- The treatment compliance estimate for abiraterone for mHSPC seems low compared to CSR data on compliance.
- The treatment compliance estimate for docetaxel for mHSPC is not applied to the same range of costs as the compliance estimate for abiraterone.
- The treatment compliance estimates for mCRPC do not take into account that they reflect discontinuations during the relevant trials. This mainly affects mCRPC treatments in the AAP+ADT arm.
- The ERG cannot find evidence that mHSPC patients who have completed their course of docetaxel and are only receiving ADT in the DOC+ADT arm have more routine bone scans than mHSPC patients in the AAP+ADT arm.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

- The submission was generally coherent and focused on the current relevant clinical evidence.
- For the economic model, the company submission uses the LATITUDE data to estimate the probabilities.
- The LATITUDE trial provides EQ-5D data, though the quality of life values estimated from this are only partially applied.
- A good range of scenario analyses are presented by the company.

1.6.2 Weaknesses and areas of uncertainty

Clinical effectiveness

- Whilst accepting that the population in the LATITUDE trial provides the best match the target patient population in the NICE scope, the company submission is weakened by being reliant upon data from only one RCT.
- There is a concern that estimates from both of the company's ITCs using NMA for the mHSPC group and the Bucher pairwise estimates for the mCRPC patients are not be robust due to the vast contextual heterogeneity

between studies. Fixed effects models had to be run due to insufficient numbers of trials and combinations of treatment arms to strengthen the networks and evidence. Had it be possible, random effects models would have been preferred.

Cost-effectiveness

The estimates of 1st line mCRPC costs and benefits may not be reliable. These are central to the cost effectiveness estimates as they provide cost offsets to the abiraterone mHSPC treatment costs. All the cost effectiveness estimates may not be reliable.

The company prefer the MSM/TA387 model over the MSM model. Due to the *ad hoc* 2.62 hazard ratio this is in large part an elaborate non-statistical method of fitting curves to the LATITUDE Kaplan Meier OS curves. The fitting of the MSM/TA387 model OS curves to the LATITUDE Kaplan Meier OS curves also seems to largely negate the reason for adopting the MSM/TA387 modelling approach.

If curves are to be fitted to the LATITUDE Kaplan Meier OS curves it may be better to use the usual well-established statistical methods, which would also allow time varying probabilities to be explored.

There may be procedural issues around using the model outputs of a previous submission as axiomatic inputs to the model of a subsequent submission. Approval of abiraterone for mCRPC prior to chemotherapy during TA387 also does not imply that the model outputs of TA387 were necessarily viewed by the Committee as reliable estimates of the most probable mCRPC OS and discontinuation curves.

The Committee for this appraisal may be more equipoise between the MSM model and the MSM/TA387 model than the company. The most important difference between them is the amount of time they model patients spending on 1st line, 2nd line and 3rd line mCRPC treatment. Alternatively, the Committee may prefer a partitioned survival analysis, or a presentation of parameterised curves that are fitted statistically to the LATITUDE rPFS and OS data by way of model validation.

There is uncertainty about what 1st line mCRPC treatments proportions should be applied subsequent to AAP+ADT, ADT and DOC+ADT treatment for mHSPC. It is also not clear whether NICE approval of abiraterone for mHSPC would over time lead to mHSPC patients receiving more than one novel agent for their metastatic prostate cancer. These proportions are likely to become more important if the models' estimates of 1st line mCRPC treatments' costs and benefits are corrected.

The company do not submit any scenario analyses that limit the extrapolation of the treatment effect, as suggested in the NICE methods guide section 5.1.16.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG make a number of revisions to the company model. The detail of these is presented in section 5.4. The main ERG revisions are:

- Applying the full LATITUDE quality of life regression so that the quality of life values reflect those observed during the trial:
- Not applying the company quality of life decrement for those who have completed a course of docetaxel for mHSPC. The ERG consider the evidence presented by the company for this as thin. There is RCT data which may suggest there is actually an increment.
- Applying a compliance estimate for mHSPC abiraterone costs based upon compliance data in the clinical study report. The company estimate derived from the LATITUDE rPFS and TTD curves seems too low, particularly towards the end of these curves.
- Equalising the frequency of bone scans for those who have completed a course of docetaxel for mHSPC with those receiving abiraterone for mHSPC in the AAP+ADT arm.

Each of these changes has a reasonable impact upon the cost effectiveness estimates.

The results summarised below take into account the abiraterone commercial access agreement but do not take into account the enzalutamide, cabazitaxel or radium-223 patient access schemes. The ERG provides a separate cPAS Appendix that takes into account the enzalutamide, cabazitaxel or radium-223 patient access schemes.

When using the MSM/TA387 model the ERG's changes taken together worsen the cost effectiveness estimates from £17,418 per QALY to £17,992 per QALY for the comparison of AAP+ADT with ADT and from £17,828 per QALY to £31,439 per QALY for the comparison of AAP+ADT with DOC+ADT.

When using the MSM model the ERG's changes taken together worsen the cost effectiveness estimates from £20,438 per QALY to £20,855 per QALY for the comparison of AAP+ADT with ADT and from £26,909 per QALY to £41,697 per QALY for the comparison of AAP+ADT with DOC+ADT.

The probabilistic estimates are aligned with these deterministic estimates.

The ERG provide a range of sensitivity and scenario analyses:

- Applying the LATITUDE Kaplan Meier data for a longer period worsens the cost effectiveness estimates.
- Assuming that DOC+ADT patients who progress have the same probability of receiving treatment for mCRPC as those in the AAP+ADT arm worsens the cost effectiveness estimate.
- Differentiating 1st line mCRPC treatments' effectiveness has little effect. However, assuming that patients prefer enzalutamide rather than abiraterone for 1st line mCRPC treatment improves the cost effectiveness estimates. Both costs and QALYs are affected due to enzalutamide not being associated with a quality of life treatment effect increment compared to ADT, whereas abiraterone is.
- Quality of life increments and decrements for ADT (post DOC+ADT) have the predictable effects.
- Not applying the LATITUDE QoL regression in full but deriving SAE and SRE decrements from values in the literature improves the cost effectiveness estimates considerably.
- Applying the company mHSPC abiraterone compliance percentage improves the cost effectiveness estimates.
- Applying the company bone scan frequencies for DOC+ADT improves the cost effectiveness estimates considerably.

2 Background

2.1 Critique of company's description of underlying health problems

The company's description of prostate cancer and newly diagnosed metastatic hormone sensitive prostate cancer (mHSPC) in terms of prevalence, symptoms and complications appears generally accurate and appropriate to the decision problem. Prostate cancer is the most common male cancer in the UK, with over 46,700 people diagnosed in 2014.[1] Approximately 18% of new cases present with metastases at first diagnosis, meaning the cancer is diagnosed too late for curative treatment to be possible as it has already spread outside the prostate gland and through the body.[1] The term metastatic hormone sensitive prostate cancer refers to people who have not received hormone therapy or have received hormone therapy but have not yet become resistant to treatment. Those with newly diagnosed mHSPC have a poorer prognosis than people who are first diagnosed with localised disease.[2, 3] Localised prostate cancer has an expected survival of at least five years after diagnosis, while only 30% of those with metastatic disease are expected to reach five-year survival.[4] The outlook for those classed as 'high-risk' at diagnosis is even worse, with life expectancy generally less than three years on conventional hormone therapy.[5-7] This is because high-risk disease is aggressive and is likely to advance more quickly.

High-risk disease is defined as having two of the following three poor prognostic factors: a Gleason score of ≥8 (describing the aggressiveness of the tumour), the presence of ≥3 lesions on a bone scan, or the presence of visceral metastases (both describing the extent of tumour spread).[4] Approximately 50% of men with newly diagnosed mHSPC are likely to have high-risk prognostic factors at diagnosis, amounting to approximately 4400 cases each year (Incidence statistics, Janssen Research & Decelopment, 2018).[1, 8] 'High-volume' is a concept previously used in mHSPC research (i.e. the CHAARTED and GETUG-AFU 15 studies) which is of similar severity to high-risk disease (three or more bone lesions and visceral metastasis) but without a specified Gleason score. As well as impacting survival, quicker progression to metastatic castrate resistant prostate cancer (mCRPC) is associated with further reduced health-related quality of life (HRQL), increased healthcare costs and greater medical resource use (MRU), affecting both patients and

the wider NHS.[9, 10] Symptoms can be highly debilitating and distressing. Over half of advanced prostate cancer patients suffer from pain, fatigue, drowsiness and bone pain. Up to 75% of people with advanced prostate cancer develop bone disease that can result in skeletal-related events (SREs) including spinal cord compression and pathological fracture,[11] both of which are associated with loss of mobility and further impaired HRQOL.[12] Patients with high-volume disease report worse HRQOL compared to men with low-volume disease as measured by the Functional Assessment of Cancer Therapy-Prostate (FACT-P) questionnaire. Diagnosis of advanced prostate cancer also carries a psychological burden. Compared with localised disease, those with advanced prostate cancer report less vitality and energy, as well as poorer social and emotional wellbeing.[13]

2.2 Critique of company's overview of current service provision

The ERG believe the company's description of current service provision for newly diagnosed mHSPC is correct.

The ultimate aims of treating newly diagnosed mHSPC are to delay disease progression (and thus extend the time to developing mCRPC), maintain HRQOL and prolong survival.[14] Prostate cancer is an androgen-dependent disease, and inhibition of testosterone is a key initial treatment strategy. Androgen deprivation therapy (ADT) has been the standard of care (SOC) in mHSPC and it is still used as monotherapy to treat 50–60% of these people in the UK.[15-17] As noted in the company submission, orchidectomy (surgical castration) and bicalutamide monotherapy are less common treatment options.(Advisory Board Report, Janssen Research & Development, 2017).[18]Although most men initially respond to ADT, the vast majority develop progressive disease within one to two years.[19] Data from the CHAARTED and STAMPEDE studies have shown that giving docetaxel (chemotherapy) in addition to ADT to men with newly diagnosed mHSPC (i.e. before they have become resistant to hormone therapy) was beneficial for health outcomes. Although unlicensed in this setting, NHS England have released a clinical commissioning policy to support the use of docetaxel with ADT in newly diagnosed mHSPC in response because of its reported survival benefits,[20, 21] and new recommendations for the use of docetaxel in addition to ADT have been implemented in most guidelines published by the urological and oncological societies.[7] Whilst ADT alone does not

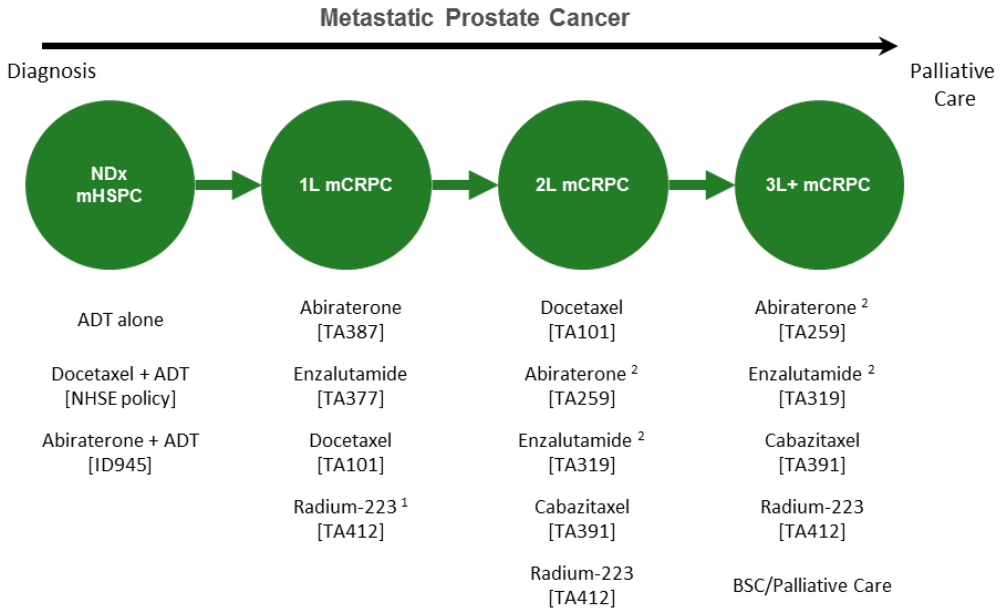
elicit comparable survival benefits, the addition of docetaxel is associated with greater toxicity and potentially severe side effects. Similarly, 20% of patients are considered clinically unsuitable for docetaxel and other psychological, social and economic factors influence an individual's suitability for treatment; such as the presence of a carer or loved one for support, proximity to chemotherapy clinics, emotional capacity to endure the toxicity of chemotherapy and religious beliefs that can prevent uptake of chemotherapy due to the alcohol content in docetaxel. As a result, some patients in the UK prefer to delay chemotherapy and would choose to receive ADT alone and, as stated in the company submission, this compromises their survival in the absence of any alternative life-prolonging therapy. Limitations of docetaxel and ADT treatment are presented in Table 1, reproduced from the Company submission, document B, Table 3 on page 17.

Table 1 Limitation of current treatment

Treatment	Limitations
Docetaxel	• Docetaxel (plus ADT) for the treatment of prostate cancer is commonly associated with numerous AEs[22] including: – Grade 4 neutropenia as well as other grade 3/4 blood and lymphatic system disorders such as anaemia, febrile neutropenia and thrombocytopenia. – Grade 3/4 gastrointestinal disorders, including nausea, diarrhoea and vomiting. – Grade 3/4 neuropathy, alopecia and fatigue • Docetaxel-associated grade 3/4 toxicities are shown to have detrimental effects on patients' QoL.[23] – One patient has described being "*unable to carry out daily chores like tidying up*" and another could "*hardly walk due to groin pain*".[24] – Docetaxel also impacts social interaction, psychological and emotional wellbeing. – The morbidity associated with docetaxel can incur significant AE costs whilst compromising the effectiveness of treatment due to resulting dose reductions and discontinuations.[25] • Docetaxel is not suitable for use in all patients, due to clinical prognostic factors (such as ECOG PS and comorbidities) as well as patient preferences[26, 27] • Docetaxel can negatively impact on carers, despite their efforts to stay positive and provide support; some have specifically

Treatment	Limitations
	mentioned the emotional impact of witnessing a family member or friend battle the disease.[24] – According to the Burden Scale for Family Caregivers Tool, 79% of caregivers for men undergoing docetaxel reported they wished they could *"run away from their current situation*", and 58% were worried about their future. • Increased use of docetaxel in mHSPC could deplete the number of chemotherapy services available for NHS patients with other cancers.
ADT alone	• Despite initial response to ADT, most patients progress to mCRPC within one to two years.[19] – Progression to mCRPC is associated with substantial burden on patients directly, and on wider society indirectly. – Patients with mCRPC have worse vitality, social functioning and mental health and more pain compared to patients with mHSPC.[9] – mCRPC is also associated with longer inpatient stays and greater number of prescriptions for outpatient drugs, all leading to increased healthcare costs.[10] • Patients with metastatic prostate cancer treated with ADT alone have life expectancy of less than four years; further reduced to less than three for patients with high-risk disease.[21, 28]
Key: ADT, androgen deprivation therapy; AE, adverse event; ECOG PS, Eastern Cooperative Oncology Group Performance Status; mCRPC, metastatic castrate resistant prostate cancer; mHSPC, metastatic hormone sensitive prostate cancer; NHS, National Health Service; OS, overall survival; QoL, quality of life.	

The care pathway for newly diagnosed metastatic disease has evolved and treatment can now be considered in terms of sequential lines of therapy, i.e. first-line treatment for mHSPC followed by a sequence of suitable regimens (first line [1L], second line [2L], etc.) for mCRPC. Several novel agents are now available and the order in which a patient may receive them is determined by prior treatment, NICE recommendation and NHS policy. The clinical pathway of care provided is reproduced from the company submission (document B, figure 4 on page 19) and presented as Figure 1. A summary of the current NICE guidelines for the treatment of metastatic prostate cancer is presented in Table 2.



Key: ADT, androgen deprivation therapy; BSC, best supportive care; mCRPC, metastatic castration-resistant prostate cancer; mHSPC, metastatic hormone-sensitive prostate cancer; NDx, newly diagnosed; NHSE, National Health Service England.
Notes: [1], If docetaxel is contraindicated or not suitable; [2], Use of abiraterone or enzalutamide in mCRPC is dependent on the prior use of docetaxel and/or prior abiraterone or enzalutamide, as per respective NICE guidance

Figure 1 Clinical pathway of care for metastatic prostate cancer in NHS England and the Company's proposed positioning for AAP

Table 2 Current NICE guidelines for the treatment of metastatic prostate cancer

Therapy	Population	Summary of NICE guidance	NICE technology appraisal or clinical guidance number
Androgen deprivation therapy (ADT)	Metastatic prostate cancer	For people who are willing to accept the adverse impact on overall survival and gynaecomastia in the hope of retaining sexual function, offer anti-androgen monotherapy with bicalutamide 150mg. Begin ADT and stop bicalutamide treatment in people who do not maintain satisfactory sexual function.	CG175[29]
Abiraterone	Castration-resistant metastatic prostate cancer previously treated with docetaxel	Abiraterone, in combination with prednisone or prednisolone, is recommended only if: • the disease has progressed on or after one docetaxel-containing chemotherapy regimen and • the manufacturer provides abiraterone in accordance with the commercial access arrangement as agreed with NHS England	TA259[30]
	Metastatic hormone-relapsed prostate cancer	Abiraterone in combination with prednisone or prednisolone is recommended, within its marketing authorisation, as an option for treating metastatic hornone-relapsed prostate cancer: • in people who have no or mild symptoms after androgen deprivation therapy has failed, and before chemotherapy is indicated • only when the company provides abiraterone in accordance with the commercial access arrangement as agreed with NHS England	TA387[31]

Docetaxel	Hormone-refractory metastatic prostate cancer	Docetaxel is recommended, within its licensed indications, as a treatment option for hormone-refractory prostate cancer only in their Karnofsky performance-status score is 60% or more. It is recommended that treatment with docetaxel should be stopped: • at the completion of planned treatment of up to 10 cycles or • if severe adverse events occur or • in the presence of progression of disease as evidenced by clinical or laboratory criteria, or by imaging studies. Repeat cycles of treatment with docetaxel are not recommended if the disease recurs after completion of the planned course of chemotherapy.	TA101[32]
Enzalutamide	Metastatic hormone-relapsed prostate cancer, before chemotherapy is indicated	Enzalutamide is recommended, within its marketing authorisation,: • in people who have no or mild symptoms after androgen deprivation therapy has failed, and before chemotherapy is indicated • when the company provides it with the discount agreed in the patient access scheme	TA377[33]
	Metastatic hormone-relapsed prostate cancer previously treated with docetaxel	Enzalutamide is recommended, within its marketing authorisation, an option for treating metastatic hormone relapsed prostate cancer in adults whose disease has progressed during or after docetaxel-containing chemotherapy, only if the manufacturer provides enzalutamide with the discount agreed in the patient access scheme.	TA316[34]
Cabazitaxel	Hormone-relapsed metastatic prostate cancer treated with docetaxel	Cabazitaxel in combination with prednisone or prednisolone is recommended in people with metastatic hormone-relapsed prostate cancer, whose disease has progressed during or after docetaxel if: • the person has an ECOG performance status of 0 or 1 • the person has had 225 mg/m^2 or more of docetaxel	TA391[35]

		• treatment with cabazitaxel is stopped when the disease progresses or after a maximum of 10 cycles (whichever happens first) In addition, cabazitaxel is recommended only if: • the company provides cabazitaxel with the discount in the patient access scheme agreed with the Department of Health, and • NHS Trusts purchase cabazitaxel in accordance with the commercial access agreement between the company and NHS England, either - pre-prepared intravenous-infusion bags, or - in vials, at a reduced price that includes a further discount reflecting the average cost of waste per patient	
Radium-223 dichloride	Hormone-relapsed prostate cancer with bone metastases	Radium-223 dichloride is recommended as an option for treating hormone-relapsed prostate cancer, symptomatic bone metastases and no known visceral metastases in adults only if: • they have had docetaxel or • docetaxel is contraindicated or is not suitable The drug is only recommended if the company provides the discount agreed in the patient access scheme	TA412[36]
Best supportive care/palliative care	Metastatic prostate cancer	Personal preferences for palliative care should be discussed as early as possible with people with metastatic prostate cancer and their partners and carers. Treatment and care plans should be tailored accordingly.	CG175[29]

3 Critique of company's definition of decision problem

3.1 Population

The NICE final scope for this appraisal specified the population as adults with newly diagnosed high risk metastatic hormone-naïve prostate cancer (mHNPC). The population addressed in the company submission is adults with newly diagnosed, high risk mHSPC. The company state that the marketing authorisation wording describes AAP as indicated for the treatment of newly diagnosed high risk mHSPC in combination with ADT. The company further state that terms mHNPC and newly diagnosed mHSPC are effectively the same because, if a patient is newly diagnosed they are, by default, hormone naïve.

3.2 Intervention

The intervention in both the NICE final scope and the company submission is abiraterone acetate (trade name Zytiga) with prednisone/prednisolone (AAP) plus androgen deprivation therapy (ADT). AAP is currently authorised in more than 100 countries worldwide for the treatment of mCRPC.[37] AAP decreases serum testosterone to undetectable levels when given with LHRH analogues.

The company provides details of abiraterone acetate in Table 2 of the submission (document B, page 11) and is reproduced by the ERG in this report as Table 3 below.

Table 3 Technology being appraised

UK approved name and brand name	**Abiraterone acetate (Zytiga®)**
Mechanism of action	Abiraterone acetate (AA) is converted *in vivo*, to abiraterone, a potent androgen biosynthesis inhibitor that selectively inhibits the enzyme 17α-hydroxylase (CYP17). CYP17 catalyses the conversion of pregnenolone and progesterone into the testosterone precursors dehydroepiandrosterone (DHEA) and androstenedion[38] CYP17 inhibition also results in increased mineralocorticoid production by the adrenal glands via a feedback loop which culminates in increased adrenocorticotropic hormone (ACTH) secretion. By inhibiting the production of both DHEA and androstenedione, AA blocks androgen biosynthesis at all sites in the body, including the testes, adrenal glands and prostatic tumour. Treatment with AA decreases serum testosterone to undetectable levels (using commercial assays) when given with LHRH agonists (or orchidectomy)[39, 40]
Marketing authorisation/CE mark status	Positive Committee for Medicinal Products for Human Use (CHMP) opinion was received on 12th October 2017. Marketing authorisation was subsequently granted on 20th November 2017.6
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Abiraterone acetate is indicated with prednisone or prednisolone for: the treatment of newly diagnosed high-risk metastatic hormone sensitive prostate cancer (mHSPC) in adults in combination with androgen deprivation therapy (ADT) the treatment of metastatic castration-resistant prostate cancer (mCRPC) in adult men who are asymptomatic or mildly symptomatic after failure of androgen deprivation therapy in whom chemotherapy is not yet clinically indicated the treatment of mCRPC in adults whose disease has progressed on or after a docetaxel-based chemotherapy regimen.[38]
Method of administration and dosage	AA is administered orally at a recommended dose of 1,000mg (two 500mg tablets) as a single daily dose in combination with 5mg prednisolone daily for mHSPC and 10mg daily for mCRPC.[38]
Additional tests or investigations	Serum transaminases should be measured prior to starting treatment, every two weeks for the first three months of treatment and monthly thereafter, until treatment discontinuation. Blood pressure, serum potassium and fluid retention should be monitored monthly. During treatment of patients with significant risk for congestive heart failure, blood pressure, serum potassium fluid retention, and other signs and symptoms of congestive heart failure should be monitored every two weeks for three months, then monthly thereafter and abnormalities corrected.
List price and average cost of a course of treatment	The NHS list price of AA 500mg tablets x 56 = £2,735.00. Treatment with AA is continued until disease progression. The median duration of treatment in men with newly diagnosed high-risk mHSPC is 24 months.[41]

	Drug cost [list price] £2,735 [28 days] Packs per year 365/28 = 13 Drug cost per patient per year* £35,652.68 *Maximum drug cost presented, assuming all patients who are initiated on abiraterone acetate stay on treatment for a full year.
Patient access scheme (if applicable)	Abiraterone acetate is available to NHS customers through a confidential Commercial Access Arrangement (CAA) with NHS England. This CAA will extend to cover the use of AAP + ADT in patients with newly diagnosed high-risk mHSPC. [redacted]
Key: AA, abiraterone acetate; CYP17, 17α-hydroxylase; DHEA, dehydroepiandrosterone; EPAR, European Public Assessment Report; LHRH, luteinising-hormone-releasing hormone; PAS, patient access scheme; SPC, summary of product characteristics.	

3.2.1 Safety

Abiraterone acetate (AA) may cause hypertension, hypokalaemia, fluid retention and cardiac failure due to increased mineralocorticoid levels. Caution is required in treating patients whose underlying medical conditions might be compromised by these contraindications (e.g. cardiac glycosides, severe renal impairment, heart failure, severe or unstable angina pectoris, recent myocardial infarction or ventricular arrhythmia).

It is recommended that potassium levels are maintained at ≥ 4.0 mM in patients with pre-existing hypokalaemia or those that develop hypokalaemia whilst being treated with AA.[38] For patients who develop Grade ≥ 3 toxicities including hypertension, hypokalaemia, oedema and other non-mineralocorticoid toxicities, treatment should be withheld and appropriate medical management should be instituted. Treatment with AA should not be reinitiated until symptoms of the toxicity have resolved to Grade 1 or baseline.[38]

For patients who develop hepatotoxicity during treatment (alanine aminotransferase [ALT] increases or aspartate aminotransferase [AST] increases above 5 times the upper limit of normal [ULN]), treatment should be withheld immediately. Re-treatment following return of liver function tests to baseline may be given at a reduced dose of 500 mg (two tablets) once daily and serum transaminases should be monitored at a minimum of every two weeks for three months and monthly thereafter. If hepatotoxicity recurs at the reduced dose of 500 mg daily, treatment should be discontinued. If patients develop severe hepatotoxicity (ALT or AST 20 times the ULN) anytime while on therapy, treatment should be discontinued and patients should not be re-treated.[38]

No dose adjustment is necessary for patients with pre-existing mild hepatic impairment but there are no data for the safety or efficacy of multiple does of AA in patients with moderate to severe hepatic impairment. It is, therefore, advised that AA is used cautiously in patients with moderate impairment and not used in patients with severe impairment.[38]

AA should be used with caution in patients with a history of cardiovascular disease and treatment should be discontinued if there is a clinically significant decrease in cardiac function. Decreased bone density may occur in people with metastatic advanced prostate cancer and the use of AA in combination with a glucocorticoid could increase this effect. Caution is also recommended in patients concomitantly treated with medicinal products known to be associated with myopathy/rhabdomyolysis. Sexual dysfunction and anaemia may occur in patients with mCRPC, including those undergoing treatment with AA.

3.2.2 Adverse reactions

The company provided details of adverse reactions observed during clinical studies and post-marketing experience in Table 1 of Appendix C, and reproduced by the ERG below. Frequency categories are defined as follows: very common (≥ 1/10); common (≥ 1/100 to < 1/10); uncommon (≥ 1/1,000 to < 1/100); rare (≥ 1/10,000 to < 1/1,000); very rare (< 1/10,000) and not known (frequency cannot be estimated from the available data). Within each frequency grouping, undesirable effects are presented in order of decreasing seriousness.

Table 4 Adverse reactions identified in clinical studies and post-marketing

System Organ Class	Adverse reaction and frequency
Infections and infestations	very common: urinary tract infection common: sepsis
Endocrine disorders	uncommon: adrenal insufficiency
Metabolism and nutrition disorders	very common: hypokalaemia common: hypertriglyceridaemia
Cardiac disorders	common: cardiac failure*, angina pectoris, atrial fibrillation, tachycardia uncommon: arrhythmia not known: myocardial infarction, QT prolongation (see sections 4.4 and 4.5)
Vascular disorders	very common: hypertension
Respiratory, thoracic and mediastinal disorders	rare: allergic alveolitis[a]
Gastrointestinal disorders	very common: diarrhoea common: dyspepsia
Hepatobiliary disorders	very common: alanine aminotransferase increased and/or aspartate aminotransferase increased [b] rare: hepatitis fulminant, acute hepatic failure
Skin and subcutaneous tissue disorders	common: rash
Musculoskeletal and connective tissue disorders	uncommon: myopathy, rhabdomyolysis
Renal and urinary disorders	common: haematuria

General disorders and administration site conditions	very common: oedema peripheral
Injury, poisoning and procedural complications	common: fractures**

* Cardiac failure also includes congestive heart failure, left ventricular dysfunction and ejection fraction decreased

** Fractures includes osteoporosis and all fractures with the exception of pathological fractures

[a] Spontaneous reports from post-marketing experience

[b] Alanine aminotransferase increased and/or aspartate aminotransferase increased includes ALT increased, AST d hepatic function abnormal.

3.3 Comparators

The NICE final scope specifies the comparators as ADT alone (including orchidectomy, luteinising hormone-releasing hormone [LHRH] agonist therapy or monotherapy with bicalutamide) and docetaxel + ADT. The comparators considered by the company differ from the NICE scope. The company state that clinical experts advised that both orchidectomy and bicalutamide monotherapy are seldom used in the UK and the company, consequently, chose to not include these comparators in their submission. The comparators presented in the company submission are ADT alone (including LHRH agonist therapy) and docetaxel + ADT. The company state that clinical experts provided validation that there is no difference in the type of ADT, thus justifying their approach. The ERG clinical expert agrees that orchidectomy and bicalutamide monotherapy are seldom used in NHS clinical practice and that it is appropriate to remove these as comparators for AAP + ADT.

3.4 Outcomes

The outcomes stated in the NICE final scope are: overall survival (OS), progression free survival (PFS), prostate specific antigen (PSA) response, adverse effects of treatment and HRQOL. The company submission includes all the outcomes listed in the NICE scope and reports additional outcomes from the LATITUDE trial: PFS following subsequent therapy (PFS2), time to symptomatic local progression, prostate cancer-specific survival, time to chronic opiate use, castration status.

3.5 Other relevant factors

The company present several factors, substantiated by UK clinical experts, that could prevent a person with newly diagnosed high-risk mHSPC from undertaking treatment

with docetaxel + ADT for reasons beyond clinical prognostic factors. These include but are not limited to:

- The presence of a carer or loved one for support, both for attending chemotherapy clinics and managing potential side effects
- Where a man lives, be it isolated or accessible by public transport to attend chemotherapy clinics, with or without a carer
- The emotional state required to endure the toxicity of chemotherapy, which is often understated
- Religious beliefs that can prevent a man from pursuing chemotherapy due to the alcohol content in docetaxel
- Being unwilling to undertake treatment

It is therefore essential that psychological, social and economic factors are considered so that clinicians and patients can make an informed judgement regarding which treatment is best suited to an individual patient.

Table 5 Comparison of NICE final scope and decision problem addressed by the company

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with newly diagnosed high risk metastatic hormone-naïve prostate cancer (mHNPC)	Adults with newly diagnosed, high-risk, mHSPC.	As per the marketing authorisation wording: AAP is indicated for the treatment of newly diagnosed high-risk mHSPC in adult men in combination with ADT. While the LATITUDE trial used the term mHNPC, this is effectively the same as newly diagnosed mHSPC because (by default) if a patient is newly diagnosed, they are hormone naïve.
Intervention	AAP + ADT	AAP + ADT	N/A
Comparator(s)	ADT alone (including orchidectomy, luteinising hormone-releasing hormone [LHRH] agonist therapy or monotherapy with bicalutamide) Docetaxel + ADT	ADT alone (including LHRH agonist therapy) Docetaxel + ADT	Orchidectomy was not included because clinical experts advised this is seldom used in the UK.(Advisory Board Report, Janssen, 2017) Bicalutamide monotherapy was not included either for the same reasons. (Advisory Board Report, Janssen, 2017) Clinical experts validated there to be no difference in the type of ADT hence justifying this approach.
Outcomes	OS PFS PSA response Adverse effects of treatment HRQL	OS PFS PSA response Adverse effects of treatment HRQL	Additional outcomes are also detailed in Table 4
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year.	Pairwise ICERs were presented against ADT alone and docetaxel + ADT	The source of evidence is different for the comparison versus ADT alone (i.e. LATITUDE head-to-head data) and the comparison versus docetaxel + ADT (i.e. Bayesian ITC) therefore cannot be combined into incremental analysis.

	The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	Adhering to the reference case, a lifetime horizon was used.	N/A
	Costs will be considered from an NHS and Personal Social Services perspective.	The reference case has been adhered to.	N/A
	The availability of any commercial access agreement for the intervention and treatments included in the economic analyses will be taken into account.	Adhering to the reference case, the CAA for AAP has been applied in all economic analysis (as detailed in Table 2)	Confidential patient access schemes which apply to relevant subsequent comparator therapies are not included in these analyses as Janssen are not privy to such information.
Key: HRQL, health-related quality of life; N/A, not applicable; NICE, National Institute for Health and Care Excellence; OS, overall survival; PFS, progression-free survival; PSA, prostate specific antigen.			

4 Clinical effectiveness

4.1 Critique of the methods of review(s)

4.1.1 Literature searching

The company submission provides full details of the searches that were undertaken to identify the included studies for the clinical effectiveness review. The major relevant databases were searched: MEDLINE. MEDLINE In-Process, EMBASE and CENTRAL for RCTs and DARE for systematic reviews. The searches were undertaken in September 2015 and updated in July 2017. The searches were restricted to reports published after 2005 and in the English language

In addition, the company searched conference proceedings from six major relevant organisations for the last four years. References of identified evidence syntheses were also scrutinised for additional publications.

The search strategies are documented in full in Appendix D and are reproducible. However, the company conducted the searches using the EMBASE.com platform, which is not accessible to the ERG. The MEDLINE and EMBASE searches combined three search facets using the Boolean operator AND: prostate cancer; abiraterone or any comparator; and RCT study design. The search of MEDLINE In-Process via Pubmed, CENTRAL and DARE excluded the study design facet, which was appropriate. The search strategies were considered fit for purpose, including both relevant controlled vocabulary and text terms with appropriate used of the Boolean operators.

The company review, however, included three company-authored reports that were published after the last search date. Following a clarification question from the ERG, the company responded that the 2018 sources are company-owned publications, which were considered relevant for inclusion, despite being outside of the pre-specified search dates, as they contain data relevant to the key outcomes reported in the LATITUDE trial[42-44]. It should be noted, however, that any relevant comparator studies published after the last search date would not have been identified.

The company submission originally included studies published prior to 2005 but in response to the ERG request for clarification, the company removed these studies from the report.

4.1.2 Inclusion criteria

The company conducted a systematic review to assess the clinical effectiveness of AAP plus ADT. The company provided details of their inclusion criteria, shown in Table 6 below. A total of 16 studies met all the inclusion criteria and were ultimately included in the company's systematic review. Of these, only the LATITUDE trial[41] was considered to match the patient population indicated in the company submission and this forms the primary evidence base of the submission. Two additional trials (CHAARTED and GETUG-AFU 15) were included in the indirect treatment comparison (with one further trial included in sensitivity analyses - STAMPEDE). Of the 16 included studies, the most commonly investigated intervention (either as intervention of interest or comparator) was conventional ADT, which was evaluated in all but five studies. Abiraterone was investigated in two studies (LATITUDE and STAMPEDE).[41, 45]

Table 6 Inclusion and exclusion criteria for the systematic review of clinical effectiveness (reproduced from Table 4, Appendix D of the company submission)

Category	Inclusion criteria	Exclusion criteria
Population	Men (aged 18 years and over) with high risk/high-volume mHSPC	Publications reporting on patient populations in the following categories: Females Children Healthy volunteers Patients with only non-cancerous prostate disease (such as benign prostatic hyperplasia) Patients with malignancies other than prostate cancer Patients with localised/locally advanced prostate cancer Metastatic prostate cancer patients who have progressed on endocrine manipulation for their disease
Interventions	Studies to be considered eligible for inclusion in the review will have reported on at least one of the following treatments: Abiraterone acetate (Zytiga®) Enzalutamide (Xtandi®) Conventional ADT drugs: Luteinising hormone-releasing hormone agonists Buserelin Histrelin Goserelin Leuprorelin Triptorelin Luteinising hormone-releasing hormone antagonists/gonadotropin releasing hormone Degarelix	Publications that do not report data specific to treatment using abiraterone acetate, ADT, docetaxel and enzalutamide

Category	Inclusion criteria	Exclusion criteria
	Anti-androgens Bicalutamide Flutamide Nilutamide Cyproterone Androgen blocker Aminoglutethimide Ketoconazole Chemotherapy Docetaxel Surgery Bilateral orchiectomy	
Comparators	No limits will be applied for comparators	N/A
Outcomes	The review will be limited to publications that report on the following outcomes: Clinical effectiveness Clinical safety	Publications that only report data on the following types of outcomes: Narrative publications, non-systematic reviews, case studies, case reports, editorials PK/PD HRQL and related PROs Cost and resource use ICERs, QALYs and other cost-effectiveness outcomes
Study type	The review will be limited to publications of studies with the following designs: RCTs	Publications of studies with the following designs: Animal studies *In vitro*/*ex vivo* studies Gene expression/protein expression studies Prospective non-randomised controlled interventional studies Prospective longitudinal observational studies

Category	Inclusion criteria	Exclusion criteria
		Retrospective longitudinal observational studies Cross-sectional studies Economic models and trial-based economic analyses Systematic reviews and meta-analyses of RCTs[a]
Time limit	Sept 2015 through to present	Studies published before 2015
Language	English language	Non-English language

Key: ADT, androgen deprivation therapy; HRQL, health-related quality of life; ICER, incremental cost-effectiveness ratio; mHNPC, metastatic hormone-naïve prostate cancer (also called castrate-sensitive, hormone-dependent, or hormone-sensitive prostate cancer); N/A, not applicable; PK/PD, pharmacokinetics/pharmacodynamics; PROs, patient reported outcomes; QALY, quality-adjusted life year; RCT, randomised controlled trial.

Notes: [a], Systematic reviews and meta-analyses of RCTs will be included and flagged. Bibliographies of these systematic reviews will be screened to check if literature searches have missed any potentially relevant studies

4.1.3 Critique of data extraction

The company state that two reviewers independently screened titles and abstracts identified by the literature searches. Secondary screening of full text articles was also independently conducted by two reviewers, although it is unclear whether these were the same reviewers who screened titles and abstracts. During the study selection any uncertainties between the two reviewers were checked by a senior reviewer. Data were extracted using a pre-specified template by one independent reviewer and validated by a second senior reviewer. The ERG consider the methods used by the company to be appropriate.

4.1.4 Quality assessment

Quality assessment was conducted for every included full text publication by the company using the National Institute of Health and Care Excellence (NICE) quality assessment tool, based on the Centre for Reviews and Dissemination (CRD) guidance.[46] The company reported the results of their quality assessment for the trials included in the indirect treatment comparison. These are presented in Table 7. The

ERG mainly agree with the company's results. The company did not provide an overall risk of bias for the STAMPEDE trial. The ERG judge this trial to be at unclear risk of bias due to the high risk scoring for performance bias.

Table 7 Summary of quality assessment for the RCTs included in the indirect treatment comparison (reproduced from Table 11, Appendix D of the company submission)

Study	Selection bias	Performance bias	Attrition bias	Detection bias	Overall risk
LATITUDE[41]	Unclear	Low risk	Low risk	Low risk	Low risk
CHAARTED[21]	High risk	High risk	Low risk	Low risk	High risk
GETUG AFU-15[28]	Low risk	Low risk	Low risk	Unclear	Low risk
STAMPEDE[45]	Low risk	High risk	Low risk	Low risk	

The ERG conducted a quality assessment of the methods used by the company for the systematic review of clinical evidence using the CRD criteria. Results are presented in Table 8.

Table 8 Quality assessment of the company's systematic review of clinical effectiveness evidence

CRD quality item	Yes/No/Unclear
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes
2. Is there evidence of a substantial effort to search for all of the relevant research?	Yes
3. Is the validity of included studies adequately assessed?	Yes
4. Are sufficient details of the individual studies presented?	Yes
5. Are the primary studies summarised appropriately?	Yes

4.1.5 Evidence synthesis

The company provide evidence for the effectiveness of AAP plus ADT from two RCTs: LATITUDE and STAMPEDE. LATITUDE[41] is a manufacturer-sponsored, multinational, randomised, double-blind, placebo-controlled Phase III trial that investigated AAP plus ADT versus ADT plus placebos (hereafter referred to as, and

considered equal to, ADT alone) in people with newly diagnosed, high-risk mHSPC. This is the only trial providing data specific to the target (i.e., licensed) population of interest, and thus is the primary evidence source for the company submission. The manufacturer-sponsored STAMPEDE study[45] represents the largest evidence base of data specific to UK clinical practice for AAP + ADT in early prostate cancer but include a broader patient population than LATITUDE and does not report data separately for HRD/HVD patients. Due to these limitations, data from the STAMPEDE trial are referenced as supportive evidence only in the company submission.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Characteristics and critique of the trials included in the systematic review of clinical effectiveness

As stated previously in section 4.1.5, the main evidence for the company submission is taken from the LATITUDE trial[41] with supporting evidence presented from the STAMPEDE trial[45]. A summary description of these two trials is presented in Table 9.

Table 9 Summary of the two RCTs presented in the review of clinical effectiveness (reproduced from Table 4, Document B of the company submission)

Study	LATITUDE (NCT01715285)[41]	STAMPEDE (NCT00268476)[45]
Study design	A manufacturer-sponsored, multinational, randomised, double-blind, placebo-controlled Phase III trial.	An investigator-sponsored, multinational (UK dominant), multi-arm multi-stage platform design incorporating a seamless Phase II/III component.
Population	Newly diagnosed high-risk mHSPC.[a] [High-risk is defined as having 2 of the following: Gleason score of ≥8, the presence of ≥3 lesions on a bone scan, or the presence of visceral metastases]	Prostate cancer that was newly diagnosed and metastatic, node-positive, or high-risk localised or disease that was previously treated with radical surgery or radiotherapy and was now relapsing with high-risk features.
Intervention(s)	AA 1,000mg daily oral dose (given once daily as four 250mg tablets) plus prednisolone 5mg daily plus ADT (LHRH agonist or orchidectomy). Selection of the LHRH agonist was at the investigator's discretion, and dosing was consistent with the respective product labelling. Patients could also have opted to undergo surgical castration in lieu of receiving ADT by LHRH analogue.	• Docetaxel + ADT 75mg/m^2 IV on Day 1 plus prednisolone 5mg BID for 21 days Q3W for a maximum of six cycles • AA 1000mg (4x 250mg) daily oral dose plus prednisolone 5mg daily plus ADT Permitted methods of ADT included bilateral orchidectomy, LHRH agonists or antagonists, dual androgen blockade, or other methods discussed with the STAMPEDE trial team. The planned duration of ADT +/- AA was 2 years in non-metastatic patients and until disease progression in metastatic patients.
Comparator(s)	ADT alone (LHRH agonist or orchidectomy). Selection of the LHRH agonist was at the investigator's discretion, and dosing was consistent with the respective product labelling. Patients could also have opted to undergo surgical castration in lieu of receiving ADT by LHRH agonist.	ADT alone. Permitted methods of ADT included bilateral orchidectomy, LHRH agonists or antagonists, dual androgen blockade, or other methods discussed with the STAMPEDE trial team. The planned duration of ADT was 2 years in non-metastatic patients and until disease progression in metastatic patients.
Supports marketing authorisation	Yes	No
Used in the economic model	Yes	Yes, for sensitivity analysis only

Study	LATITUDE (NCT01715285)[41]	STAMPEDE (NCT00268476)[45]
Rationale for use/non-use in the model	Pivotal trial supporting this indication.	Provides supportive randomised data of the benefits of AAP + ADT; however, this is not specific to the population of interest in this submission.
Reported outcomes specified in the decision problem	• OS (co-primary endpoint) • rPFS (co-primary endpoint) • Time to next SRE[b] • Time to PSA progression (Prostate Cancer Working Group 2 criteria) • Time to subsequent therapy for prostate cancer • Time to initiation of chemotherapy • Time to pain progression • Safety, including time to treatment discontinuation • HRQL, including BPI-SF, FACT-P, BFI and **EQ-5D-5L**	• OS (primary endpoint) • FFS (intermediate primary endpoint) • Safety • Symptomatic skeletal events • PFS • PSA-specific survival • HRQL, including EQ-5D and EORTC QLQ-C30 with the prostate-specific module QLQ PR25[c]
All other reported outcomes	• PSA response rate • PFS following subsequent therapy (PFS2) • Time to symptomatic local progression • Prostate cancer-specific survival • Time to chronic opiate use • Castration status	N/A

Key: AA, abiraterone acetate; ADT, androgen deprivation therapy; BFI, Brief Fatigue Inventory; BID, twice daily; BPI-SF, Brief Pain Inventory-Short Form; EQ-5D-5L, EuroQoL; FACT-P, Functional Assessment of Cancer Therapy – Prostate; FFS, failure-free survival; HRQL, health-related quality of life; IV, intravenous; LHRH, luteinising hormone releasing hormone; mHSPC, metastatic hormone sensitive prostate cancer; N/A, not applicable; OS, overall survival; PFS, progression-free survival; PSA, prostate specific antigen; rPFS, radiographic progression-free survival; SRE, skeletal-related event; Q3W, every 3 weeks.
Notes:[a], Patients could have received up to 3 months treatment with ADT prior to randomisation; [b], economic model uses SRE rates; [c], HRQL data have not yet been published from the STAMPEDE study.

The baseline demographics and disease characteristics were well-balanced across treatment groups in the LATITUDE trial and are shown in Table 10. The majority of patients (>95%) had a Gleason score ≥8 and ≥3 bone lesions (96% in the AAP plus ADT group, 95% in the ADT alone group). Post-hoc analysis showed that 487 patients (82%) in the AAP plus ADT group and 468 patients (78%) in the ADT alone group had 'high-volume' disease, defined as the presence of visceral metastases or ≥4 bone lesions with ≥1 beyond the vertebral bodies and pelvis (as per CHAARTED and GETUG-AFU 15 studies, discussed further in section 4.3). The extent of disease was similar between groups, as was median PSA level (25ng/mL in the AAP plus ADT group and 23ng/mL in the ADT alone group), demonstrating that patients with high-risk and high-volume disease are closely comparable.

There was comparable distribution in the use of hormonal therapy, surgery or radiotherapy across treatment groups. Most patients received prior hormonal therapy, comprising predominantly of a gonadotrophin-releasing hormone (GnRH) analogue (75%) and first generation anti-androgens (62%). A smaller percentage of patients had undergone an orchidectomy (12%). Although this is higher than what is usually seen in UK clinical practice, the company state there is no clinical difference between orchidectomy and LHRH, and the form of ADT would not impact the effect of AAP.

Table 10 Baseline characteristics of the LATITUDE intention to treat population (reproduced from Table 6, Document B of the company submission)

	AAP + ADT (n=597)	**ADT Alone (n=602)**
Age, median years (range)	68 (38–89)	67 (33–92)
Median PSA level before ADT, ng/mL (range)	25.4 (0–8,775.9)	23.1 (0.1–8,889.6)
ECOG PS, n (%)	0: 326 (54.6) 1: 245 (41.0) 2: 26 (4.4)	0: 331 (55.0) 1: 255 (42.4) 2: 16 (2.7)
Gleason score at initial diagnosis, n (%)	<7: 4 (0.7) 7: 9 (2) ≥8: 584 (98)	<7: 1 (0.2) 7: 15 (2) ≥8: 586 (97)
Baseline pain score (BPI-SF Item 3), n (%)	N: 570 0–1: 284 (50) 2–3: 123 (22) ≥4: 163 (29)	N: 579 0–1: 288 (50) 2–3: 137 (24) ≥4: 154 (27)

	AAP + ADT (n=597)	ADT Alone (n=602)
≥3 bone metastases at screening, n (%)	586 (98.2)	585 (97.2)
High-risk at screening, n (%)	597 (100)	601 (100)
Gleason score ≥8 + ≥3 bone lesions	573 (96)	569 (95)
Gleason score ≥8 + measurable visceral disease	82 (14)	87 (14)
≥3 bone lesions + measurable visceral disease	84 (14)	85 (14)
Gleason score ≥8 + ≥3 bone lesions + measurable visceral disease	71 (12)	70 (12)
Extent of disease, n (%)	596 (100)	600 (100)
Bone	580 (97)	585 (98)
Liver	32 (5)	30 (5)
Lungs	73 (12)	72 (12)
Node	283 (47)	287 (48)
Prostate mass	151 (25)	154 (26)
Viscera	18 (3)	13 (2)
Soft Tissue	9 (2)	15 (3)
Other	2 (0.3)	0
Bone lesions at screening, n (%)		
0	6 (1.0)	7 (1.2)
1–2	5 (0.8)	10 (1.7)
3–10	202 (33.8)	208 (34.6)
11–20	109 (18.3)	97 (16.1)
>20	275 (46.1)	280 (46.5)
Previous prostate cancer therapy, n (%)	560 (94)	560 (93)
Radiotherapy	19 (3)	26 (4)
Hormonal	559 (96)	558 (93)
GnRH agonists/antagonists[a]	449 (75)	450 (75)
Orchidectomy[a]	73 (12)	71 (12)
First-generation androgen receptor agonists	373 (62)	371 (62)
Other	7 (1)	10 (2)
Time from GnRH agonist/antagonist to first dose of study drug, median months (range)	1.08 (0.1–3.0)	1.08 (0.1–3.5)
[Post-hoc] High-volume disease, n (%)	487 (81.5)	468 (77.7)

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; BPI-SF, Brief Pain Inventory – Short Form; ECOG, Eastern Cooperative Oncology Group; GnRH, gonadotropin-releasing hormone; ITT, intention-to-treat; PS, performance status; PSA, prostate specific antigen.
Notes: [a], within 3 months prior to randomisation.
Source: Fizazi et al. 2017[41] LATITUDE CSR, 2017 Fizazi et al. 2018[44]

Table 11 presents a summary of results for the key outcomes for disease progression and overall survival for both LATITUDE and STAMPEDE.

Table 11 Summary of co-primary endpoints of the trials included in the systematic review of clinical effectiveness (reproduced from Table 9, Document B of the company submission)

Study	**LATITUDE** [ITT]		**STAMPEDE** [Arm G vs. Arm A]		**STAMPEDE** [Arm G vs. Arm C]	
Treatment	AAP + ADT	ADT alone [a]	AAP + ADT	ADT alone [a]	AAP + ADT	Docetaxel + ADT
ITT	597	602	960	957	377	189
Metastatic (%)	597 (100)	602 (100)	500 (52.1)	502 (52.5)	227 (60.2)	115 (60.8)
Patient population	NDx high-risk mHSPC		mHSPC			
Data cut	31-Oct-16		10-Feb-17		04-Mar-17	
Median follow-up	30.4 months		40 months		48 months	
Progression-free survival						
	Radiographic PFS		**PFS** [b]			
Events (%)	239 (40.0)	354 (58.8)	173 (34.6)	301 (60.0)	94 (41.4)	62 (53.9)
Median	33	14.8	-	-	-	-
[95% CI]	29.57-NE	14.69-18.27	-	-	-	-
HR	0.47		0.43		0.69	
[95% CI]	0.39-0.55		0.36-0.52		0.50-0.95	
p-value	<0.0001		-		0.02	
Failure-free Survival [c]						
HR	-	-	0.31		0.56	
[95% CI]	-	-	0.26-0.37		0.42-0.75	
p-value	-	-	-		<0.001	
Overall Survival						
Events (%)	169 (28.3)	237 (39.4)	150 (30.0)	218 (43.4)	89 (39.2)	38 (33.0)
Median	NE	34.7	-	-	-	-
[95% CI]	NR-NR	33.05-NR	-	-	-	-
HR	0.62		0.61		1.13	
[95% CI]	[0.51-0.76]		0.49-0.75		0.77-1.66	
p-value	<0.0001		0.195×10^{-7}		0.53	

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; CI, confidence interval; DOT, duration of treatment; HR, hazard ratio; IQR, inter-quartile range; ITT, intent to treat; m1, metastatic; mHSPC, metastatic hormone sensitive prostate cancer; NDx, newly diagnosed; NR, not reached; OS, overall survival; rPFS, radiographic progression-free survival; PSA, prostate-specific androgen; SRE, skeletal-related events; Tx, treatment.
Notes: [a], Placebos + ADT; [b], PFS defined as radiologic or clinical progression or death from prostate cancer, [c], FFS defined as radiologic, clinical, PSA progression or death from prostate cancer
Source: Fizazi et al. 2017[41] LATITUDE CSR, 2017; James et al. 2017 [45]Sydes et al. 2017[47] Rydzewska et al. 2017[48]

Progression-free survival

For progression-free survival (PFS), treatment with AAP + ADT significantly delayed disease progression in patients with newly diagnosed high-risk mHSPC when compared with ADT alone in the LATITUDE trial. Treatment with AAP + ADT resulted in a 53% reduction in the risk of radiographic progression or death (HR=0.47 [95% CI: 0.39–0.55]; p<0.001). At three years, 47% of patients in the AAP + ADT arm remained event-free, compared to only 21% of those in the ADT alone arm.

While STAMPEDE did not consider PFS, the company argue that the outcome failure free survival (FFS) is considered to be generally comparable by the clinical community. In the metastatic (M1) subgroup of STAMPEDE, treatment with AAP + ADT was associated with a 69% reduction in the risk of biochemical failure, progression or death compared with ADT alone (HR=0.31 [95%CI: 0.26–0.37]; p<0.0001). Although the M1 subgroup is broader than the licensed indication for abiraterone, the company state that results for this comparable endpoint of FFS provide strong supporting evidence for the benefit of AAP + ADT over ADT in prolonging time to disease progression. Whilst the post-hoc analysis from STAMPEDE comparing AAP + ADT with docetaxel + ADT was not pre-specified, and thus not statistically powered to detect clinical differences in treatment, results for the M1 subgroup showed treatment with AAP + ADT was associated with a 44% reduction in the risk of biochemical failure, progression or death (HR=0.56 [95% CI: 0.42–0.75]; p<0.001)

Overall survival

In the LATITUDE trial, treatment with AAP + ADT was associated with a 38% reduction in the risk of death compared with ADT alone (HR=0.62 [95%CI: 0.51–0.76]; p<0.001).7 The overall survival rate at three years was 66% in the AAP plus ADT group and 49% in the ADT alone group. There was an imbalance in the proportion of patients who received life-extending subsequent therapies (20.9% in the AAP plus ADT arm vs. 40.9% in the ADT alone arm), which could result in the standard ITT analysis of OS underestimating the true OS benefit for AAP. Therefore, additional pre-specified OS analysis using the IPCW methodology was conducted to adjust for patients who switched to other therapies. This analysis showed AAP plus ADT significantly improved survival compared to ADT alone, with an improved

HR=0.48 (95% CI: 0.36–0.63; p<0.0001). Results from STAMPEDE are consistent with these results. Treatment with AAP + ADT was associated with a 39% reduction in the risk of death compared to ADT alone (HR= 0.61 [95% CI: 0.49–0.75]; p<0.0001).

Health-related quality of life

Patients' responses to the Visual Analogue Scale (VAS) and their health utility scores were significantly improved (p<0.05) when treated with AAP + ADT in LATITUDE and time to HRQL degradation was significantly by 4 to 6 months (15%), as measured by the Functional Assessment of Cancer Therapy – Prostate (FACT-P) total score, as well as consistently delaying worsening of pain-related symptoms by 24%. The worsening of physical wellbeing on treatment with AAP + ADT was also delayed by 25%, allowing patients to experience a longer time before their physical condition got worse. Significant improvements, as measured by the Brief Fatigue Inventory (BFI) in fatigue were also observed with AAP + ADT treatment. Median time to pain progression, measured by the BPI short form, was not reached for patients who received AAP + ADT and was 16.6 months for patients who received ADT alone, demonstrating a significant delay until pain progression (HR=0.70 [95% CI: 0.583–0.829], p<0.0001). These data indicate a 31% reduction in the risk of pain progression. The 36-month event-free rate was 55.5% for AAP + ADT versus 37.9% for ADT alone.

Secondary endpoints

Table 12 presents the summary of secondary endpoints for the LATITUDE trial.

Table 12 Summary of secondary endpoints for the LATITUDE intention to treat population (reproduced from Table 11, Document B of the company submission)

	AAP + ADT (n=597)	ADT alone (n=602)
Time to pain progression		
Events, n (%)	233 (39.0)	289 (48.0)
Median months (95% CI)	NR (36.5, NR)	16.6 (11.1, 24.0)
HR (95% CI) [p-value]	0.70 (0.58–0.83) [<0.001]	
Time to subsequent prostate cancer therapy		
Events, n (%)	191 (32.0)	322 (53.5)
Median months (95% CI)	NR ([redacted])	21.6 ([redacted])
HR (95% CI) [p-value]	0.42 (0.35–0.50) [<0.001]	
Time to life-extending subsequent therapy for prostate cancer		
Events, n (%)	125 (20.9)	246 (40.9)
Median months (95% CI)	[redacted]	[redacted]
HR (95% CI) [p-value]	[redacted]	
Time to initiation of chemotherapy		
Events, n (%)	109 (18.3)	191 (31.7)
Median months (95% CI)	NR ([redacted])	38.9 ([redacted])
HR (95% CI) [p-value]	0.44 (0.35–0.56) [<0.001]	
Time to PSA progression		
Events, n (%)	241 (40.4)	434 (72.1)
Median months (95% CI)	33.2 (27.6, NR)	7.4 (7.2, 9.2)
HR (95% CI) [p-value]	0.30 (0.26–0.35) [<0.001]	
Time to next SRE		
Events, n (%)	[redacted]	[redacted]
Median months (95% CI)	NR (NR, NR)	NR (NR,NR)
HR (95% CI) [p-value]	0.70 (0.54–0.92) [0.009]	

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; CI, confidence interval; HR, hazard ratio; ITT, intention-to-treat; NR, not reached; PSA, prostate specific antigen; SRE, skeletal-related event.
Source: Fizazi et al. 2017[41] LATITUDE CSR, 2017 European Public Assessment Report[37].

In the LATITUDE trial, treatment with AAP+ADT significantly reduced the time to subsequent therapy for prostate cancer. The median time to subsequent therapy was not reached in the AAP + ADT group, it was 21.6 months for the ADT group (HR=0.415 [95%CI: 0.346–0.497], p<0.0001). Twice as many ADT alone patients required life-extending subsequent therapy (either docetaxel, enzalutamide, cabazitaxel, radium-233 or AAP) compared with those who received AAP+ADT (40.9% versus 20.9% respectively). The median time to life-extending subsequent therapy was not reached in the AAP + ADT group and was 29.5 months in the ADT

group (HR=0.37 [95%CI: 0.29–0.45]; p<0.0001). Of those who received life-extending subsequent therapy at any time, docetaxel was the most common treatment after AAP+ADT or ADT alone (17.8% and 31.1%, respectively).

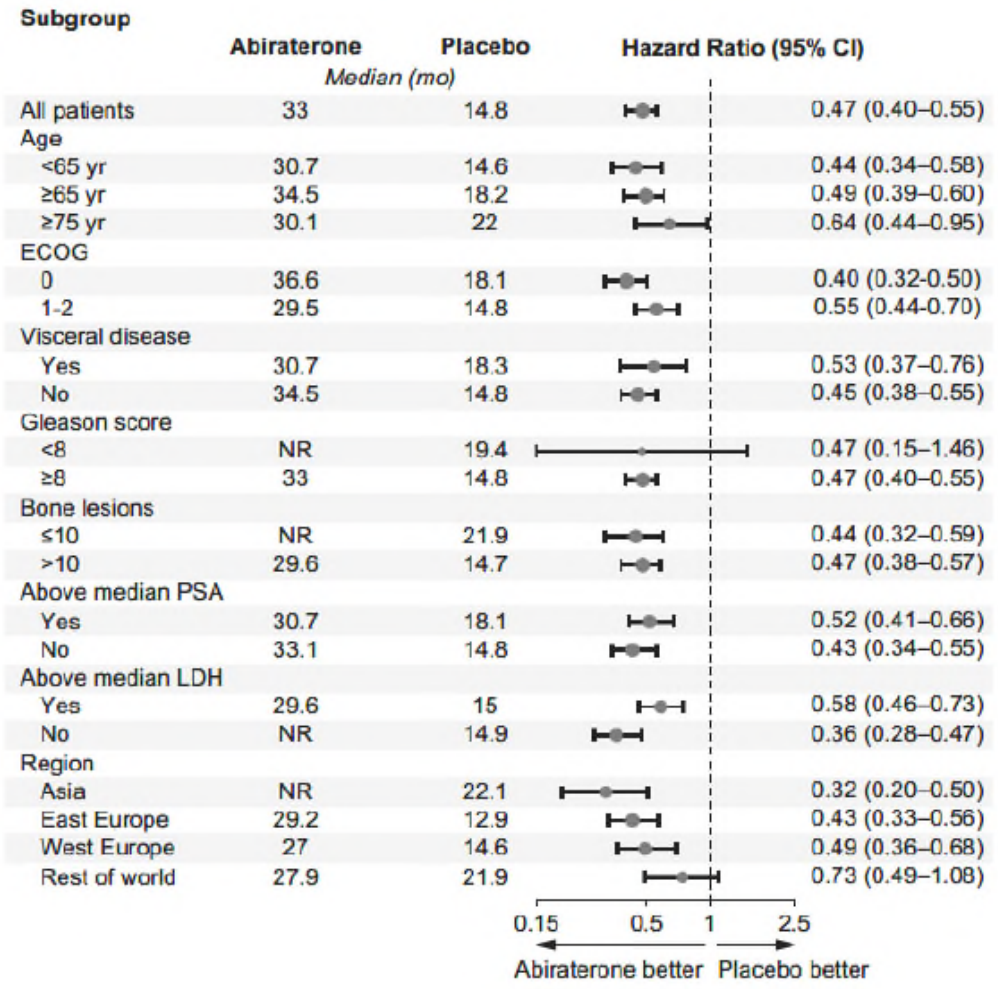
The median time to initiation of chemotherapy was not reached in the AAP + ADT group and 38.9 months in the ADT group. This translated to a 56% reduction in risk for initiating chemotherapy.

Time to PSA progression in the LATITUDE trial was defined as a 25% increase in PSA from baseline, and an increase in absolute value of 2ng/mL or more, after 12 weeks of treatment. The median time to PSA progression was 33 months in the AAP + ADT arm compared to 7 months in the ADT alone arm (HR=0.30 [95% CI: 0.26–0.35]; p<0.0001).

Treatment with AAP + ADT significantly reduced the risk of SREs by 30% (HR=0.70 [95%CI: 0.539,0.916], p=0.0086), although median time to SRE was not reported in either arm. However, it should be noted that this analysis was based on data for a small number of events, and the results should be interpreted with caution.

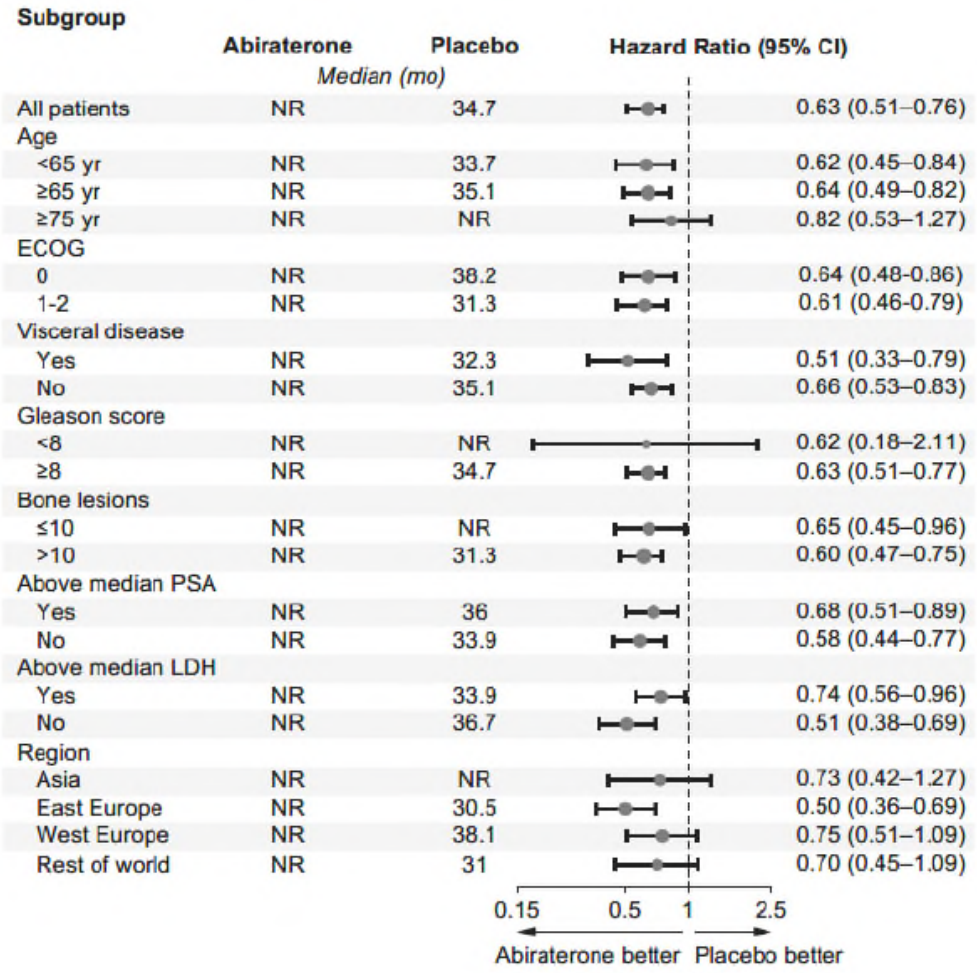
Deaths due to prostate cancer occurred less frequently in the AAP + ADT group compared to the ADT alone group (20.4% vs. 32.3%, respectively). This resulted in a statistically significant improvement in prostate cancer-specific survival for the AAP + ADT group compared to the ADT alone group (HR=0.55 [95% CI: 0.44, 0.69]; p<0.0001)

LATITUDE subgroup analyses are presented in Figures 2 and 3.



Key: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; ITT, intention-to-treat; LDH, lactate dehydrogenase; mo, months; NR, not reached; PSA, prostate specific antigen; rPFS, radiographic progression-free survival.
Source: Fizazi et al. 2017[41]

Figure 2 Subgroup analysis of radiographic progression free survival from the LATITUDE intention to treat population (reproduced from Figure 21, Document B of the company submission)



Key: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; ITT, intention-to-treat; LDH, lactate dehydrogenase; mo, months; NR, not reached; OS, overall survival; PSA, prostate specific antigen.
Source: Fizazi et al. 2017[41]

Figure 3 Subgroup analyses of overall survival from the LATITUDE intention to treat population (reproduced from Figure 22, Document B of the company submission)

Details of a meta-analyses of LATITUDE and STAMPEDE overall survival and disease progression data, which was independently conducted by Rydzewska et al.,[48] are presented in Table 13. Results of these analyses show a significant survival benefit of AAP + ADT versus ADT alone (HR=0.62 [95%CI: 0.53–0.71]; p=0.55 x 10-[10]) and a consistently significant benefit of AAP + ADT versus ADT was demonstrated for disease progression (HR=0.45 [95%CI: 0.40–0.51]; p=0.66 x10-[36]). The company note that STAMPEDE M1 subgroup is broader than the licensed indication for

abiraterone in mHSPC, these data should only be considered as strong supporting evidence of AAP's clinical effectiveness, with the direct evidence from LATITUDE being the most appropriate source to inform the economic modelling in the licensed population.

Table 13 Meta-analyses of AAP+ADT versus ADT alone for the outcomes overall survival and disease progression (reproduced from Table 16, Document B of the company submission)

OS			
	Direct Evidence: AAP+ADT vs. ADT	**Direct Evidence: AAP+ADT vs. ADT**	**Meta-Analysis: AAP+ADT vs. ADT**
Trial Name	LATITUDE	STAMPEDE	Rydzewska et al. 2017[48]
Population	ITT	M1	
HR [95% CrI]	0.62 [0.51, 0.76]	0.61 [0.49, 0.75]	0.62 [0.53, 0.71]
Disease progression (i.e. rPFS or PFS)[a]			
	Direct Evidence: AAP+ADT vs. ADT	**Direct Evidence: AAP+ADT vs. ADT**	**Meta-Analysis: AAP+ADT vs. ADT**
Trial Name	LATITUDE	STAMPEDE	Rydzewska et al. 2017[48]
Population	ITT	M1	
HR [95% CrI]	0.47 [0.39, 0.55]	0.43 [0.36, 0.52]	0.45 [0.40, 0.51]
Key: AAP, abiraterone acetate prednisone/prednisolone; ADT, androgen deprivation therapy; CI, confidence interval; CrI, credible interval; HR, hazard ratio; ITT, intent to treat; M1, metastatic; OS, overall survival; rPFS, radiographic progression-free survival. **Notes:** [a], PFS=radiological or clinical progression-free survival			

Adverse reactions

Treatment exposure

The median treatment duration in the safety population of the LATITUDE trial was 24 months in the AAP + ADT arm and 14 months in the ADT alone arm.[41] [redacted] A total of 91.8% of patients in the AAP + ADT group and 86.0% of patients in the ADT alone group received ≥6 cycles of study drug; 54.5% and 29.7% of patients, respectively, received ≥24 cycles.

Dose reductions were reported for [redacted]% of patients treated with AAP + ADT and [redacted]% of patients treated with ADT alone, while dose interruptions were reported for

[redacted]% and [redacted]% of patients, respectively. A [redacted] percentage of patients in the AAP + ADT group had dose interruptions of prednisolone due to AEs compared with the ADT group ([redacted]% vs. [redacted]%). A [redacted] percentage of patients in each treatment group ([redacted]% AAP + ADT and [redacted]% ADT alone) had additional prednisolone prescribed by the investigator for more than two weeks to manage drug-related toxicity pertaining to insufficient control of mineralocorticoid effects.
The company report that data on treatment duration reported in STAMPEDE are not comparable to data reported in LATITUDE.

Summary safety data

A summary of treatment emergent adverse events (TEAEs) is presented in Tables 14 and 15. TEAEs were reported by a higher number of people in the AAP+ADT group than for ADT alone. The most frequently reported TEAEs in the LATITUDE trial (preferred terms reported in ≥20% of patients) in either the AAP + ADT or ADT alone arm were hypertension (37% versus 22%, respectively), hypokalaemia (20% versus 4%) and back pain (18% versus 20%). Commonly reported SAEs (≥1% of patients in either the AAP + ADT or ADT alone group) included pneumonia (1.8% versus 0.3%, respectively), spinal cord compression (1.7% versus 1.8%) and urinary retention (1.5% versus 1.7%). The most frequently reported AEs leading to treatment discontinuation (reported in ≥1% of patients in either the AAP + ADT or ADT alone group) were spinal cord compression (0.8% versus 1.0% of patients, respectively) and bone pain (0.5% versus 1.0%, respectively). Cases of discontinuation for hypokalaemia, hypertension and cardiac disorders were rare. A post-hoc analysis of safety data for LATITUDE patients with HVD was consistent with the intention to treat population. This post hoc group had a similar baseline characteristics profile to those of the ITT.

Table 14 Summary of adverse reactions in the LATITUDE safety population (reproduced from Table 13, Document B of the company submission)

	LATITUDE	
	AAP + ADT (n=597)	**ADT alone (n=602)**
Any TEAE, n (%)	558 (93.5)	557 (92.5)
Drug-related	336 (56.3)	269 (44.7)
Any serious TEAE, n (%)	165 (27.6)	146 (24.3)
Drug-related	29 (4.9)	12 (2.0)
Grade 3–4 TEAE, n (%)	374 (62.6)	287 (47.7)
Drug-related	162 (27.1)	67 (11.1)
Discontinuation due to TEAE, n (%)	73 (12.0)	61 (10.1)
Drug-related	21 (3.5)	11 (1.8)
Death due to TEAE, n (%)	28 (4.7)	24 (4.0)
Drug-related	3 (0.5)	3 (0.5)

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; TEAE, treatment-emergent adverse event.
Source: Fizazi et al. 2017[41] European Public Assessment Report[37]

Table 15 Treatment emergent Grade 3-4 adverse events reported in at least 1% of patients in the LATITUDE safety population (reproduced from Table 14, Document B of the company submission)

	AAP + ADT (n=597)			ADT Alone (n=602)		
	Total	**Grade 3**	**Grade 4**	**Total**	**Grade 3**	**Grade 4**
Any TEAE, n (%)	374 (62.6)	342 (57.3)	32 (5.4)	287 (47.7)	265 (44.0)	22 (3.7)
Vascular disorders	127 (21.3)	126 (21.1)	1 (0.2)	65 (10.8)	64 (10.6)	1 (0.2)
Hypertension	121 (20.3)	121 (20.3)	0	60 (10.0)	59 (9.8)	1 (0.2)
Cardiac disorder						
Any	74 (12)	15 (3)	5 (1)	47 (8)	6 (1)	0
Atrial fibrillation	8 (1)	2 (<1)	0	2 (<1)	1 (<1)	0
Metabolism and nutrition disorders	98 (16.4)	90 (15.1)	8 (1.3)	42 (7.0)	39 (6.5)	3 (0.5)
Hypokalaemia	62 (10.4)	57 (9.5)	5 (0.8)	8 (1.3)	7 (1.2)	1 (0.2)
Hyperglycaemia	27 (4.5)	26 (4.4)	1 (0.2)	18 (3.0)	18 (3.0)	0
Hyperkalaemia	7 (1.2)	5 (0.8)	2 (0.3)	9 (1.5)	9 (1.5)	0
Investigations	69 (11.6)	62 (10.4)	7 (1.2)	47 (7.8)	45 (7.5)	2 (0.3)
ALT increase	33 (5.5)	31 (5.2)	2 (0.3)	8 (1.3)	8 (1.3)	0
AST increase	26 (4.4)	25 (4.2)	1 (0.2)	9 (1.5)	9 (1.5)	0
LDH increase	11 (1.8)	10 (1.7)	1 (0.2)	9 (1.5)	9 (1.5)	0
Weight increase	6 (1.0)	6 (1.0)	0	6 (1.0)	6 (1.0)	0
Musculoskeletal and connective tissue disorders	55 (9.2)	55 (9.2)	0	72 (12.0)	72 (12.0)	0
Bone pain	20 (3.4)	20 (3.4)	0	17 (2.8)	17 (2.8)	0
Back pain	14 (2.3)	14 (2.3)	0	19 (3.2)	19 (3.2)	0
Pain in extremity	7 (1.2)	7 (1.2)	0	12 (2.0)	12 (2.0)	0

	AAP + ADT (n=597)			ADT Alone (n=602)		
	Total	**Grade 3**	**Grade 4**	**Total**	**Grade 3**	**Grade 4**
Arthralgia	6 (1.0)	6 (1.0)	0	15 (2.5)	15 (2.5)	0
Musculoskeletal pain	4 (0.7)	4 (0.7)	0	6 (1.0)	6 (1.0)	0
Muscular weakness	3 (0.5)	3 (0.5)	0	7 (1.2)	7 (1.2)	0
Nervous system disorders	35 (5.9)	32 (5.4)	3 (0.5)	35 (5.8)	31 (5.1)	4 (0.7)
Spinal cord compression	12 (2.0)	12 (2.0)	0	10 (1.7)	7 (1.2)	3 (0.5)
Infections and infestations	31 (5.2)	29 (4.9)	2 (0.3)	19 (3.2)	17 (2.8)	2 (0.3)
Pneumonia	10 (1.7)	9 (1.5)	1 (0.2)	3 (0.5)	3 (0.5)	0
Urinary tract infection	6 (1.0)	6 (1.0)	0	5 (0.8)	5 (0.8)	0
Renal and urinary disorders	30 (5.0)	29 (4.9)	1 (0.2)	29 (4.8)	28 (4.7)	1 (0.2)
Urinary retention	10 (1.7)	10 (1.7)	0	8 (1.3)	8 (1.3)	0
Haematuria	6 (1.0)	6 (1.0)	0	3 (0.5)	3 (0.5)	0
Blood and lymphatic system disorders	26 (4.4)	21 (3.5)	5 (0.8)	35 (5.8)	33 (5.5)	2 (0.3)
Anaemia	15 (2.5)	12 (2.0)	3 (0.5)	27 (4.5)	26 (4.3)	1 (0.2)
General disorders and administration site conditions	26 (4.4)	26 (4.4)	0	39 (6.5)	37 (6.1)	2 (0.3)
Fatigue	10 (1.7)	10 (1.7)	0	14 (2.3)	14 (2.3)	0
Asthenia	4 (0.7)	4 (0.7)	0	7 (1.2)	7 (1.2)	0
General physical health deterioration	4 (0.7)	4 (0.7)	0	6 (1.0)	6 (1.0)	0

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; AE, adverse event; ALT, alanine aminotransferase; AST, aspartate transaminase; LDH, lactate dehydrogenase; TEAE, treatment-emergent adverse events.
Source: Fizazi et al. 2017[41] European Public Assessment Report[37]

No new safety signals were identified in the LATITUDE trial compared to those already characterised through the use of AAP in mCRPC, has across the two established licensed indications. AAP + ADT was well tolerated, with a comparable incidence of TEAEs to ADT alone. In line with its known safety profile, the most frequently reported Grade 3 or 4 TEAEs were mineralocorticoid-associated AEs.7 However, all events were medically manageable, only rarely required treatment discontinuation and seldom led to serious consequences. The safety results from LATITUDE are further supported by the STAMPEDE trial which also demonstrated that AAP + ADT was well tolerated, with a comparable incidence of Grade 3 to 5 AEs to ADT alone in patients with metastatic and non-metastatic prostate cancer. Table 16 presents a summary of Grade 3-4 adverse events reported in STAMPEDE.

Neutropenia and febrile neutropenia were more frequent after treatment with docetaxel (13% and 17%, respectively), compared to AAP (1% for both outcomes). Hypertension and hypokalaemia were reported more frequently by patients treated with AAP + ADT, compared to that observed in the LATITUDE trial. Treatment discontinuation due to AEs with AAP + ADT was similar in LATITUDE and STAMPEDE, and was also comparable with respect to ADT alone (10%). Of note, data on the occurrence of AEs by pre-specified metastatic subgroups were not provided

Table 16 Grade 3-5 adverse events reported in the STAMPEDE safety population (reproduced from Table 15, Document B of the company submission)

	AAP + ADT vs. ADT alone		AAP + ADT vs. docetaxel + ADT	
	AAP + ADT (n=948)	**ADT alone (n=960)**	**AAP + ADT (n=373)**	**ADT + Doc (n=172)**
AE, n (%)				
Endocrine disorders	129 (14)	133 (14)	49 (13)	15 (9)
Febrile neutropenia	-	-	3 (1)	29 (17)
Neutropenia	-	-	4 (1)	22 (13)
Cardiovascular disorders	92 (10)	41 (4)	32 (9)	6 (3)
Hypertension	44 (5)	13 (1)	-	-
MI	10 (1)	9 (1)	-	-
Cardiac dysrhythmia	14 (1)	2 (<1)	-	-
Musculoskeletal disorders	68 (7)	46 (5)	33 (9)	9 (5)
Gastrointestinal disorders	49 (5)	40 (4)	28 (8)	9 (5)
Hepatic disorders	70 (7)	12 (1)	32 (9)	1 (1)
Increased ALT levels	53 (6)	4 (<1)	-	-
Increased AST levels	10 (1)	2 (<1)	-	-
General disorders	45 (5)	29 (3)	21 (6)	18 (10)
Fatigue	21 (2)	15 (2)	-	-
Oedema	5 (1)	0	-	-
Respiratory disorders	44 (5)	23 (2)	11 (3)	12 (7)
Dyspnoea	18 (2)	7 (1)	-	-
Laboratory abnormalities	34 (4)	21 (2)	11 (3)	9 (5)
Hypokalaemia	12 (1)	3 (<1)	-	-

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; Doc, docetaxel; MI, myocardial infarction.
Source: James et al. 2017[45]; Sydes et al. 2017[47]

4.2.2 Critique of statistical techniques used in trial

4.3 Critique of trials identified and included in the indirect comparison and/ or multiple treatment comparison

The company presented results of an indirect treatment comparison (ITC). The company reported their criteria for considering whether the trials included in the

systematic review of effectiveness were eligible for inclusion in the ITC. The criteria were reported were that the trials:

- Contributed data to the ITC of AAP + ADT versus ADT + docetaxel
- Reported comparable outcomes of interest
- Were sufficiently comparable with regards to study design, treatment and patient-level characteristics.

A total of three of the 16 trials (LATITUDE, CHAARTED and GETUG-AFU 15)[21, 28, 41] were included in the global base case network and one additional trial (STAMPEDE) was included in sensitivity analyses. Both CHAARTED and GETUG-AFU 15 were phase 3, open-label RCTs. As described earlier, LATITUDE was a phase 3, double-blind RCT and STAMPEDE was a multi-arm, multi-stage phase 2/3 trial. The company state that trials that did not report data separately for HRD/HVD populations were excluded from the ITC. The STAMPEDE trial was, therefore, excluded as the trial did not report data separately for HRD/HVD patients, but was included in sensitivity analyses due to the clinical importance of this large scale trial. The ERG agree that the STAMPEDE trial does not provide sufficiently comparable data for the considered patient population to be included in the ITC.

The four trials included in the ITC were linked in a network via a standard ADT arm based on the assumption that the ADT/standard of care arms were all similar. The company state that clinical opinion confirmed that differences in docetaxel administration would not have a significant impact on outcomes and the company, therefore, determined that the docetaxel arms of the trials were similar. Details of the interventions evaluated by the trials included in the ITC are presented in Table 17. The company state that the population enrolled in the LATITIDE trial is closest to the HVD *de novo* population considered in the company submission. All patients in the LATITIUDE trial had HRD determined by patients having at least two of the following: Gleason score $\geq$8; presence of $\geq$3 lesions on a bone scan; presence of measurable visceral (excluding lymph node disease) metastasis.

Table 18 presents baseline demographics and disease characteristics of participants from the RCTs included in the ITC. In general, participant and disease characteristics

were fairly well balanced with the exception of prostate specific antigen (PSA) level before ADT, Gleason score and Eastern Cooperative Oncology Group (ECOG) performance status (the STAMPEDE trial reported World Health Organisation [WHO] performance status instead of ECOG).

Table 17 Summary characteristics of the interventions evaluated in the trials included in the company's indirect treatment comparison (reproduced from Tables 5 and 6, Appendix D of company's submission)

Name of trial	Intervention	Comparator	Population abbreviation	Population description
CHAARTED[21]	ADT: LHRH receptor agonist or an LHRH receptor antagonist or orchidectomy; anti-androgens were given at the investigators' decision.	ADT + docetaxel: Docetaxel was given as 75mg/m^2 every 3 weeks for a maximum of six cycles.	ITT HVD NDx HVD	Patients with HVD as well as those with LVD, and patients with NDx disease as well as those with prior local treatments Patients with NDx metastatic HVD AND patients who had received prior local treatments Patients with NDx metastatic HVD
GETUG-AFU 15[28]	ADT: LHRH receptor agonist alone or combined with non-steroidal anti-androgens, or orchiectomy	ADT + docetaxel: Docetaxel was given as 75mg/m^2 every 3 weeks for a maximum of nine cycles.	ITT HVD NDx HVD	Patients with HVD as well as those with LVD, and patients with NDx disease as well as those with prior local treatments Patients with NDx metastatic HVD AND patients who had received prior local treatments Patients with NDx metastatic HVD
LATITUDE[41]	LHRH or surgical castration + placebo	AAP + ADT: AA was given as 1,000mg daily (once daily as four 250mg tablets), while prednisolone was given as 5mg daily.	NDx HRD ITT NDx HVD&HRD	Patients with NDx disease; all patients have HRD Patients with NDx HVD and HRD
STAMPEDE[45]	SoC: Hormone therapy for at least 2 years with gonadotropin-releasing hormone agonists or antagonists or, only between 2006 and 2011 for patients with non-metastatic disease, oral anti-androgens alone. Orchiectomy was an allowable alternative to drug therapy. Patients received orchiectomy, LHRH-based therapy, or bicalutamide (anti-androgen)	SoC + docetaxel: Docetaxel was given as 75mg/m^2 every 3 weeks for a maximum of six cycles. AAP + SoC: AA was given as 1,000mg daily (once daily as four 250mg tablets), while prednisolone was given as 5mg daily.	M1	Patients with NDx metastatic disease; HRD or HVD status of patients is unknown

Key: AA, abiraterone acetate; ADT, androgen deprivation therapy; AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; LHRH, luteinising hormone releasing hormone; SoC, standard of care. HRD, high risk disease; HVD, high-volume disease; ITT, intent-to-treat; LVD, low volume disease; M1, metastatic; NDx, newly diagnosed

Table 18 Baseline characteristics of the participants of the RCTs included in the company's indirect treatment comparison (reproduced from Table 8, Appendix D of the company's submission)

Study	Treatment	Sample size, n	Baseline characteristics				
			Age, median years	PSA level before ADT, median (range)	ECOG PS, n (%)	Gleason score at diagnosis, n (%)	Metastases at diagnosis, n (%)
LATITUDE[41]	AAP + ADT	597	68	• 25.4	• 0: 326 (54.6) • 1: 245 (41.0) • 2: 26 (4.4)	<7: 4 (0.7) 7: 9 (2) ≥8: 584 (98)	597 (100)
	ADT alone	602	67	• 23.1	• 0: 331 (55.0) • 1: 255 (42.4) • 2: 16 (2.7)	<7: 1 (0.2) 7: 15 (2) ≥8: 586 (97)	602 (100)
CHAARTED[21]	ADT + Doc	397	64	50.9	0: 277 (69.8) 1: 114 (28.7) 2: 6 (1.5)	4–6: 21 (5.3) 7: 96 (24.2) 8–10: 241 (60.7) Unknown: 39 (9.8)	Low: 134 (33.8) High: 263 (66.2)[a]
	ADT alone	393	63	52.1	0: 272 (69.2) 1: 115 (29.3) 2: 6 (1.5)	4–6: 21 (5.3) 7: 83 (21.1) 8–10: 243 (61.8) Unknown: 46 (11.7)	Low: 143 (36.4) High: 66 (16.8)[a]
GETUG AFU-15[28]	ADT + Doc	192	63	26.7	0: 181 (99) 1–2: 2 (1)	<7: 84 (45) ≥8: 103 (55)	128 (67)
	ADT alone	193	64	25.8	0: 176 (96) 1–2: 7 (4)	<7: 78 (41) ≥8: 113 (59)	144 (76)
STAMPEDE[45]	AAP + ADT	960	67	51	0: 745 (78) 1/2: 215 (22)[b]	≤7: 221 (23) 8–10: 715 (74) Unknown: 24 (2)	500 (53)

Study	Treatment	Sample size, n	Baseline characteristics				
			Age, median years	PSA level before ADT, median (range)	ECOG PS, n (%)	Gleason score at diagnosis, n (%)	Metastases at diagnosis, n (%)
	ADT alone	957	67	56	0: 744 (78) 1/2: 213 (22)[b]	≤7: 223 (23) 8–10: 721 (75) Unknown: 13 (1)	502 (53)
	AAP + ADT	377	67	56	0: 79%[b]	Not Reported	60%
	ADT + Doc	189				Not Reported	
	ADT + Doc	592	65	70	0: 461 (78) 1+: 131 (22)[b]	≤7: 110 (19) 8–10: 436 (74) Unknown: 46 (8)	362 (61)
	ADT alone	1,184	65	67	0: 922 (78) 1+: 262 (22)[b]	≤7: 282 (24) 8–10: 810 (68) Unknown: 92 (8)	724 (61)

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; Doc, docetaxel; ECOG, Eastern Cooperative Oncology Group; MAMS, multi arm multi stage; PS, performance status PSA, prostate specific antigen.

Notes: a, volume of metastases; b, WHO performance status.

Outcome definitions differed across the trials. GETUG-AFU 15 used Response Evaluation Criteria for Solid Tumors (RECIST) version 1.0, while LATITUDE used RECISIT 1.1 definitions for radiographic progression-free survival (rPFS). Progression-free survival (PFS), failure-free survival (FFS), time to clinical progression and time to castration resistant prostate cancer (CRPC) were only available for one trial each. All trials reported overall survival (OS) and used similar definitions. Summary details of the primary outcomes are presented in Table 19.

Table 19 Summary of primary outcomes reported by the RCTs included in the company's indirect treatment comparison (reproduced from Table 9, Appendix D of the company's submission)

Outcome	CHAARTED	GETUG-AFU 15	LATITUDE	STAMPEDE
OS	Time between randomisation and death from any cause	Time between randomisation and death from any cause	Time between randomisation and death from any cause	Time between randomisation and death from any cause
rPFS	Not reported	Time from randomisation to the occurrence of radiographic progression or death from any cause (based on RECIST 1.0)	Time from randomisation to the occurrence of radiographic progression or death from any cause (based on PCWG2 and RECIST 1.1)	Not reported
FFS	Not reported	Not reported	Not reported	Time to first evidence of at least one of: • Biochemical failure • Progression either locally, in lymph nodes, or in distant metastases • Death from prostate cancer
Time to CRPC (biochemical, symptomatic or radiographic)	Time to documented clinical or serologic progression with a testosterone level of less than 50ng per decilitre	Not reported	Not reported	Not reported

Outcome	CHAARTED	GETUG-AFU 15	LATITUDE	STAMPEDE
Time to clinical progression (symptomatic or radiographic)	Time from randomisation to: • Increasing symptoms of bone metastases • Progression according to the RECIST 1.0 • Clinical deterioration due to cancer according to the investigator's opinion	Not reported	Not reported	Not reported
Key: CRPC, castration-resistant prostate cancer; FFS, failure-free survival; OS, overall survival; PCWG2, Prostate Cancer Working Group 2; PFS, progression-free survival; RECIST, Response Evaluation Criteria in Solid Tumors; rPFS, radiographic progression-free survival.				

The proportions of people receiving subsequent treatment in the included trials are presented in Table 20.

Table 20 Proportions of people receiving subsequent treatment after relapse/progression

Treatment	LATITUDE[a]		STAMPEDE[b]		CHAARTED[c]		GETUG-AFU 15[d]	
	ADT + Placebo (n=469)	ADT + AAP (n=314)	ADT (n=535)	ADT + AAP (n=248)	ADT (n=287)	ADT + Docetaxel (n=238)	ADT (n=149)	ADT + Docetaxel *
Abiraterone acetate	53 (11%)	10 (3%)	120 (22%)	8 (3%)			36 (24%)	33
Cabazitaxel	30 (6%)	11 (4%)	28 (5%)	15 (6%)	37 (13%)	57 (24%)	15 (10%)	16
Docetaxel	187 (40%)	106 (34%)	200 (37%)	115 (46%)	137 (48%)	54 (23%)	127 (85%)	
Enzalutamide	76 (16%)	30 (10%)	138 (26%)	25 (10%)			12 (8%)	15
Radium-223	27 (6%)	11 (4%)	24 (4%)	19 (8%)				
Abiraterone and/or enzalutamide					104 (36%)	105 (44%)		

a. Data from Fizazi 2017[41]– Percentages are calculated from the numbers of people who discontinued treatment and were eligible for subsequent therapy
b. Data from James 2017[45] – percentages calculated from the numbers with progression
c. Data from Sweeney 2015[21] – percentages are calculated from those with serological progression/clinical progression. Numbers for clinical progression only are ADT 228 and ADT+D 180.
d. Data from Gravis 2016[28] – * the paper reports 27/149 treated for progressive disease in the ADT arm. Unclear how many patients were treated for progression in the ADT+D arm.

Statistical comparison of AAP versus docetaxel (DOC) for the primary outcomes OS and rPFS was only possible using ITC methods. The patient populations of two RTCs, CHAARTED (790 participants) and GETUG-AFU 15 (385 participants), which compared DOC +ADT with ADT alone using post-hoc selected sub-groups of newly diagnosed patients with high volume disease (HVD), were considered to be comparable with those in LATITUDE. The company used Bayesian network meta-analyses with fixed effects to find the indirect results of AAP+ADT versus DOC+ADT. The results suggest non-significant effects for OS (HR 0.92, 95% CrL 0.69-1.23) and rPFS (HR 0.76, 95% CrL 0.53-1.10) presented in Table 21 but with Bayesian probabilities of 71.8% and 92.9%, respectively, suggesting AAP+ADT is a better life prolonging treatment option. Various sensitivity analyses examined the effect of post-hoc selection of the HVD patients rather than the high risk disease (HRD) group of LATITUDE; the inclusion of the M1 group from STAMPEDE (for both AAP+ADT and DOC+ADT) and the inclusion of those treated prior to current treatment or not. The results of the sensitivity analyses varied but there was a consistent trend in favour of AAP+ADT.

Results of sensitivity analyses of time to skeletal-related events (SRE) were similar in the indirect comparison between AAP+ADT and DOC+ADT, [redacted] but with a Bayesian pairwise probability of [redacted].

Only two RCTs, LATITUDE (AAP+ADT versus ADT) and GETUG-AFU 15 (DOC+ADT versus ADT, presumably newly diagnosed HVD patients) could be included into an ITC for the assessment of secondary outcome measures of safety. No sensitivity analyses were reported. When the AAP+ADT group (n=597) was indirectly compared to the DOC+ADT group (n=189), [redacted] However, AAP+ADT was found [redacted]

Table 21 Base case results and sensitivity analyses results of Bayesian ITC (Synthesised from Tables 18 and 19 from Document B, pages 80-81)

Outcome for original and various sensitivity analyses		**AAP + ADT vs. ADT alone**			**ADT alone vs. docetaxel + ADT**			**AAP + ADT vs. dox + ADT**	**ITC**	
		LATITUDE		**STAMPEDE**	**CHAARTED**	**GETUG-AFU 15**	**STAMPEDE**	**STAMPEDE**	**AAP + ADT vs. docetaxel + ADT**	
		ITT	**HV *post-hoc***	**M1**	**NDx HV**	**NDx HV**	**M1**	**M1**	**HR (95% CrI)**	**P_{AA-Doc}**
OS (95 % CI)	MAIN	**x**			**X**	**x**			**0.92 (0.69,**	**71.8**
	sa	x		x	X	x	x	x	0.91 (0.76,	84.5
	sa		x		X	x			0.85 (0.63,	86.7
	sa		x	x	X	x	x	x	[redacted]	[redacted]
	sa	x			x[d]	x[d]			[redacted]	[redacted]
	sa								[redacted]	[redacted]
rPFS (95% CI)	MAIN	**x**			**X**	**x**			**0.76 (0.53,**	**92.9**
	sa	x			x[d]	x[d]			0.71 (0.49, 1.02)	96.8 %

Key: AAP, abiraterone acetate + prednisolone; ADT, androgen deprivation therapy; dox, docetaxel; HR, hazard ratio; HV, high-volume; ITT, intent-to-treat; M1, metastatic disease; NDx, newly diagnosed; OS, overall survival; rPFS, radiographic progression-free survival; SA, sensitivity analysis.

Notes: $PA_{A>Doc}$, Bayesian pairwise probability for ADT+AAP being more effective compared with ADT+DOC; [a], Definitions of rPFS differed across trials; [b], Time to CRPC data; [c], FFS data, [d] included prior treated

The Functional Assessment of Cancer Therapy-Prostate (FACT-P) and the Brief Pain Inventory (BPI) quality of life measures, looked at differences of change from baseline for both AAP+ADT and DOC+ADT treatment groups over four time points 3, 6, 9 and 12 months from the LATITUDE (ITT) and CHAARTED (HVD) studies. Here another variation of sub-group analyses were conducted whereby HRD together with HVD patients in LATITUDE were selected post-hoc. At 3 months, AAP+ADT had a significant positive and beneficial increase on FACT-P over DOC+ADT, with difference of change being 4.20 (95% CrL 1.18-7.19) and the probability of the AAP patients having better quality of life to DOC being 99.7%. AAP estimates improved further over time as did the DOC estimates but not to the same extent and never to the level of AAP, although differences between AAP and DOC were not significant by 6 months or even at 1 year. BPI results showed larger decreases in pain estimates for indirect comparisons between AAP+ADT and DOC+ADT, but the results were not significant. Pain in the DOC+ADT group increased with time whereas with

AAP+ADT initially improved and then remained steady if not further reduced. The sensitivity analyses were comparable for FACT-P and BPI.

With regard to the effectiveness of AAP+ADT compared with other treatments including DOC+ADT, further indirect comparisons were conducted by the company for people with disease progression to metastatic castration resistant prostate cancer (mCRPC) although these were not presented in the clinical effectiveness section of the submission. Again, no direct head-to-head trial comparing abiratone to docetaxel was identified indicating the need for an ITC. On this occasion the company used Buchers pairwise comparisons using four trials: COU_AA_302, which compared abiratone with prednisolone as placebo; TAX327, which compared docetaxel with a mitoxantrone as placebo; ALSYMPCA, which compared radium-223 with prednisolone as placebo; and PREVAIL, which compared enzalutamide with prednisolone as placebo. Thus, each compared a treatment (AAP, radium-223, enzalutamide and docetaxel) to a 'similar' control and assessed OS. For rPFS the other co-primary outcome, only three trials could be connected, COU-AA-302, ALSYMPCA and PREVAIL. The submission focuses on the positive AAP results and the high Bayesian probabilities, see Table 22 below for the ERG replication of results [redacted].

4.4 Critique of the indirect comparison and/ or multiple treatment comparison

Abiratone compared with docetaxel for the treatment of mHSPC or mHNPC patients

With no direct head-to-head comparison of abiratone to docetaxel available the ERG agree that this gap could be bridged using and Indirect Treatment Comparison (ITC) and that the Bayesian Network Meta-Analysis (NMA) was appropriate. The company may have considered doing a Matching-Adjusted Indirect Comparisons (MAIC) instead where those in each study who fulfil the required target population and baseline characteristics are matched. However, this approach requires having the data for individuals from al the included studies and the matching often means that many of the observations are not comparable and are dropped and the results lack robustness because of poor sample size. NMA is a reasonable option.

The ERG replicated the NMA results using WinBUGS14 (50,000 burn-in and 100,000 iterations) with reference to examples and programs from NICE DSU Technical Support

Document 2.[49] The company submission used 50,000 run-in iteration phase and a 50,000 iteration phase for parameter estimation. Throughout, the company have use fixed effect models. Random effect models may have been preferred and conducted where possible by the ERG but most were not resolvable probably due to the limited number of studies.

Replication of results presented in Table 18 Document B for OS and rPFS
OS: the trials compared were LATITUDE ITT (30.4 mo), CHAARTED (newly diagnosed HVD sub-group 53.7 mo) and GETUG-AFU 15 (newly diagnosed HVD subgroup 43 mo). The ERG considered the fixed effects as per company model but also attempted a random effects model. The programmes for each may be found in DSU Document 2 Example 7a and 7b. Our findings show the fixed effects to be similar to the company submission (OS between AAP+ADT versus docetaxel+ADT, HR= 0.920, 95% CrL 0.689-1.22). The random effects model resulted in an HR of 0.894, 95% CrL: 0.258, 2.979, so slightly more benefit to the AAP+ADT but very wide credible limits (CrL). The other various sub-group analyses were also replicated using a fixed effects model. Depending on the groups used the estimate varied between 0.63 up to 1.23, which is between all the credible limits.

rPFS: the trials with relevant data for this outcome were LATITUDE ITT and GETUG-AFU 15 (newly diagnosed HVD subgroup) and again the ERG replicated ITC results [fixed effects model] for the assessment of AAP+ADT versus DOC+ADT were similar to those reported in the company submission (HR= 0.770, 95% CrL: 0.538-1.11). **Note:** random effects model did not resolve, which is to be expected with just 2 studies for 3 treatments, and thus too many parameters to estimate.

Verification of secondary outcome comparisons between trials and other treatments
Several safety and HRQL measures were compared across the trials and treatments. The ERG performed ‘trial comparisons’ for all of these, using the program in DSU Example 3b:
For the safety measures, comparisons between the HRs were performed using ‘trial arms’ and the same programs as for OS and rPRS above but on the OR’s rather than HR’s. This was because the binary data over time within each group used by the company were not provided. Nevertheless, the company estimates were comparable to those of the ERG (see Table 22 below).

or the HRQL measures, the company took an 'arm comparison' approach; however, the company submission only gave differences of mean changes and relevant CIs. Arm comparisons require actual mean changes from baseline for each arm in each trial, along with their relevant precision measure - but such information was not available. A referenced paper only had the same summary[a] estimates (without CIs) but again did not provide mean changes. As a result, while the point HR estimates are similar for the trial comparison approach, some CrLs differ to those in the submission.

Table 22 Safety and HRQL results - Bayesian ITC (Reconstruction of Tables 20 and 21; pages 83-86; Document B)

Saftey: LAITUDE and GetUG-AFU 15					**HRQL§: LAITUDE and CHAARTERED**			
Trial comparison					**Trial rather than Arm comparison**			
AAP+ADT vs ADT alone	versus	DOC+ADP vs ADT alone			AAP+ADT vs ADT alone	versus	DOC+ADP vs ADT alone	
Recalculation of Table 20 in Document B					**Recalculation of Table 21 in Document B**			
OR	**HR**	**CrL 2.5%**	**CrL 97.5%**		**Differences***	**HR**	**CrL 2.5%**	**CrL 97.5%**
Anaemia	0.065	0.036	0.118		**FACT-P** 3mo	4.196	[redacted]	[redacted]
Hot Flush	3.763	2.216	6.400		6 mo	2.487	[redacted]	[redacted]
AST	0.529	0.263	1.067		9 mo	3.067	-0.112	6.250
Constipation	0.158	0.068	0.372		12mo	2.347	-0.877	5.576
ALT	0.606	0.317	1.162		**BPI** 3 mo	-0.150[1]	[redacted]	[redacted]
Odema	0.144	0.063	0.330		6 mo	-0.761	[redacted]	[redacted]
					9 mo	-0.851	[redacted]	[redacted]
					12 mo	-0.451	[redacted]	[redacted]

§ The company also performed analysis for the HRQL measure using non ITT patients in LAITUDE who were classed as the post hoc HVD group also reported for the CHAARTED trial. The ERG did not repeat here since there was no valid reason why this considered useful here was not for other outcomes.

[a] NOTE: The reference by Feyerabend[43] shows BPI at 3 months has a difference of the mean changes for CHAARTED HVD as 44, in line with the others in that column, but in the submission this was presented as -.01.

Validity of the NMA approach for the comparison abiratone versus docetaxel

There were a number of key differences between the different trials incorporated into the NMAs above. These were:

- Differing target patient groups making ITT comparisons impossible, with respect to:
  - Being newly diagnose, and /or primary progressive
  - High Volume and High Risk – The company make some attempt to justify these are the same –but even within LATITUDE, while there were commonalities, the sub-groups did not entirely match
- Variable ADT doses in the control arm and with different definitions
- Variable docetaxel doses
- Different patterns of subsequent therapies during follow-up
- Varying previous therapies (recall some were not newly diagnosed)
- Reporting variations
- Different definitions of how to measure rPFS
- Length of studies

Since the company felt they had to compare DOC with AAP the resulting estimates are of interest albeit with huge reservations for taking them forward into the economic modelling given the degree of clinical heterogeneity highlighted above. In addition to the conceptual heterogeneity, no account of statistical heterogeneity, consistency or fit were reported in the main submission documents. If inconsistent, the results for the same treatment combinations via different routes will differ to another. Some fit statistics were provided, but not consistently nor commented on. These limitations could impact on the economic modelling and such estimates will require caution and various scenarios to reflect these concerns.

Comparison of abiratone with other treatments for patients with disease progression

The evidence of progression into mCRPC (castration resistant) is given in Appendix Q of the company submission. The company recognised that the observed heterogeneity between trials was not ideal (Appendix Q5 page 137). In particular, there was some notable clinical heterogeneity between trials including a range of differences in patient baseline characteristics. These include:

- Controls being different (although company suggest these are comparable!)
- Follow up times (i.e., survival or progression) differing from 21 to 36 months

- Some trials adjust for treatment switching (an inevitable problem) and different methods are used for this adjustment, IPE and IPCW
- Differing definitions of rPFS
- Various previous therapies
- Baseline characteristics differing between trials (only the below are admitted to)
    - Populations targeted
    - Levels of PSA

Despite these inconsistencies, comparisons of AAP+ADT versus other drugs for this patient group were needed to carry forward into any cost analyses. Rather than gaining more evidence, the company attempted indirect estimates. This time the company considered a different method to NMA (used for the mHNPC/mHSPC patient group); they chose Bucher pairwise estimates which are simple to perform and easy to understand and known to work best on "triangle structures" as is the case here. What these do not allow is for better efficiency by each trial control group being 'pooled'. The ERG understand why this approach was chosen; the company were anxious about doing a complete network analysis because of the above heterogeneity issues. The ERG agree with the company's choice. In addition, attempts to run complete NMA models by the ERG did not converge. The suggested pairwise separate comparisons was probably the only viable option even though it does not address the heterogeneity concerns highlighted above and brings issues of robustness in to question. The result was eight separate combinations. The ERG have replicated these and confirm that they are as the submission suggest (page 136 Appendix Q Table 56) but would like to reiterate that they cannot be thought of as anything but as indicators and not as robust estimates. The results are given in Table 23 below.

The interpretation of the results given by the company requires some attention.
OS: the company state that AAP+ADT has slightly lower risks, if the adjustments for treatment switching are applied. Given that not all the trials adopt a treatment switching adjustment, this strategy has to be questioned – unless there are good reasons for treatment switching to be more valid in these trials over the others (why should some be adjusted and other not). Taking the results as they stand, [redacted] suggesting that AAP will be at least as equivalent to DOC – the company however only reflect on the Bayesian probabilities. As a precautionary the ERG suggest that ongoing

economic models be based on scenarios reflecting the credible intervals around these estimates.

Table 23 ITC between AAP and other treatment in the mCRPC patient group (The ERG replicated the Bucher estimates based on Table 55 leading to Table 56 of the Appendices document)

			Each trial results			ITC results		
			HR	Low	Upper	HR	LCL	UCL
OS	AAP+Radium	COU-AA-302	0.806	0.697	0.931	[redacted]		
		ALSYMPCA	0.745	0.562	0.987			
ipe -treat swtich	AAP+Radium*	COU-AA-302	0.741	0.6	0.882	[redacted]		
		ALSYMPCA	0.745	0.562	0.987			
	AAP+Enz	COU-AA-302	0.806	0.697	0.931	[redacted]		
		PREVAIL	0.77	0.67	0.88			
ipe -treat switch AND ipcw -treat switch	AAP+ENZ **	COU-AA-302	0.741	0.6	0.882	[redacted]		
		PREVAIL	0.66	0.57	0.77			
	APP+doc	COU-AA-302	0.806	0.697	0.931	[redacted]		
		TAX327	0.76	0.62	0.94			
ipe -treat switch	AAP+Doc*	COU-AA-302	0.741	0.6	0.882	[redacted]		
		TAX327	0.76	0.62	0.94			
rPFS	AAP vs. PP Radium 223 vs. placebo	COU-AA-302	0.52	0.45	0.61	[redacted]		
Def[n] different		ALSYMPCA	0.64	0.54	0.77			
	AAP vs. PP Enzalutamide vs. placebo	COU-AA-302	0.52	0.45	0.61	[redacted]		
		PREVAIL	0.19	0.15	0.23			

rPFS: The company do not seem to fully interpret the results in their submission, focusing only on positive AAP results and the high Bayesian probabilities. [redacted]

As for OS, scenarios reflecting the credible intervals are advocated for any further economic modelling.

Overall, the company conclude that abiratone to be at least equivalent to other treatments based on these analyses, on a sensitivity analysis including STAMPEDE data and on two

previous not truly comparable systematic reviews[50, 51] for both the mHSPC and the mCRPC patient popultions. The ERG would agree this to be fair provided further claims are not made.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG was largely able to verify the company's NMA results for the mHSPC patient group using either the programs supplied in Appendix D1 pages 29-31 or comparable programs from NICE Decision Support Unit (DSU) TSD 2[49] when pertinent data were not available. Similarly the ERG confirmed the ITC results using the Bucher's approach.

4.6 Conclusions of the clinical effectiveness section

The ERG are satisfied that the methods used to conduct the systematic review of clinical effectiveness are appropriate.

The submission presents results from the LATITUDE study providing evidence of the benefits of AAP over ADT for the treatment of men with mHSPC. The benefit found in LATITUDE is evident for the primary outcomes of overall survival and progression measured by rPFS and extends to the secondary outcomes for safety and quality of life. The results of LATITUDE are similar to those from the STAMPEDE study. However, the STAMPEDE patient group was broader and while the company have conducted similar analyses on a post hoc subgroup meant to be similar to the LATITUDE population, they rightly have not combined them in any further analyses.

Less reliable are the company results of AAP compared to other treatments, predominately docetaxel. With no head-to-head studies available, these were compared using indirect methods. The company chose NMA at this stage, which the ERG agree, was sensible. When conducting the NMA the company used the recommended WinBUGS program from the NICE DSU TSD 2.[49] They were restricted to only fixed effects models because of the lack of studies and links between treatment groups. Further concerns are the many aspects of heterogeneity between the studies, all recognised by the company. So while the ERG confirm the results provided showing abiraterone to be at least equivalent to docetaxel, there is a concern that estimates from these results will not be robust. There were no checks of statistical heterogeneity or consistency commented on. As such any economic modelling on these estimates will require caution and various scenarios to reflect these concerns.

The company also attempted to assess the use of AAP+ADT for patient with disease progression (mCRPC) again compared with other subsequent treatments. Here they concentrate on docetaxel, radium-223 and enzalutamide. The more robust method of NMA was not conducted and instead the company used Bucher pairwise comparisons. While NMA are more useful when making choices between multiple alternatives, the ERG confirm that NMA models did not converge probably due to the limited number of studies and data so that Bucher estimates were a reasonable alternative. For this patient group too, the estimates show abiraterone to be comparable with other treatments. However, since checks of statistical heterogeneity or fit were not provided and as before the conceptual heterogeneity (e.g., differences in study populations, study setting, follow-up procedures, outcome measures) were extensive caution for further economic modelling is warranted.

5 COST EFFECTIVENESS

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

5.1.1 State objective of cost effectiveness review. Provide description of manufacturers search strategy and comment on whether the search strategy was appropriate. If the manufacturer did not perform a systematic review, was this appropriate?

Reports of cost effectiveness were sought by the company by searching MEDLINE AND EMBASE (via Embase.com), MEDLINE In-Process (vis Pubmed), NHS Economics Evaluation Database (NHS EED) and HTA Database (via Cochrane Library) and Econlit (via Ebsco) in September 2015 and updated in July 2017. The searches were restricted to studies published between 2005 and 2017.

The search strategies are documented in full in Appendix G and are reproducible however the company conducted the MEDLINE and EMBASE searches using the EMBASE.com platform which is not accessible to the ERG.

The MEDLINE and EMBASE searches combined three search facets using the Boolean operator AND: prostate cancer; abiraterone or comparator; and economic/cost terms.
The search strategies were considered fit for purpose, including both relevant controlled vocabulary and text terms with appropriate used of the Boolean operators. However, the ERG identified errors which were clarified by the company as documentation errors:
Date ranges: Table 17 (Embase and MEDLINE) imposed date range 2005-2015 while Table 19 (NHS EED and HTA Database) was restricted to 2015-2017. The company confirmed that all searches were run initially in 2015 with a start date of 2005, and then updated in 2017. The company removed pre-2005 studies, which had initially been included in the review of cost-effectiveness.

Table 18: The company confirmed that the heading should have been MEDLINE In process (via Pubmed).

Modifications to final set:Errors were identified by the ERG in Table 18 (lines 5-6) and Table 19 (lines 5-8). The company provided the corrected search strategies.

5.1.2 State the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

Inclusion criteria for the cost-effectiveness review are shown in Table 24

Table 24 Inclusion and exclusion criteria for the systematic review of cost-effectiveness, reproduced from Table 16, Appendix G of the company submission

Category	Inclusion criteria	Exclusion criteria	Exclusion code
Population	Men (aged 18 years and over) with mHSPC	Publications reporting on patient populations in the following categories • Females • Children • Healthy volunteers • Non-cancerous prostate disease (such as benign prostatic hyperplasia) • Cancer other than prostate • Localised/locally advanced prostate cancer patients • Metastatic prostate cancer patients who have progressed on endocrine manipulation for their disease	Population not of interest
Interventions	Abiraterone acetate, ADT, docetaxel and enzalutamide	Publications that do not report data specific to treatment using abiraterone acetate, ADT, docetaxel and enzalutamide	Intervention not of interest
Comparisons	No restriction based on treatment comparisons reported/not reported	N/A	N/A
Outcomes	The review will be limited to publications that report on the following outcomes: • Direct costs • Indirect costs • Other healthcare resource use • ICERs, QALYs, and other cost- effectiveness outcomes	Publications that only report data on the following types of outcomes: • Pharmacokinetics/pharmacodynamics • Clinical efficacy • Clinical safety • HRQL and related PROs • Epidemiological outcomes	Relevant outcomes unreported
Date	2005–2017, inclusive	Publications published before 2005	Date
Duplicate	N/A	Publications that are duplicates of other publications in the search yield	Duplicate
Publication types	N/A	Publications of the following types: • Narrative publications • Non-systematic reviews • Case studies • Case reports Editorials	Publication type not of interest
Other criteria	Only English language articles/conference abstracts will be included	Journal articles and conference abstracts without English full-text	Non-English

Key: ADT, androgen deprivation therapy; HRQL, health-related quality of life; ICER, incremental cost-effectiveness ratio; mHSPC, metastatic hormone-sensitive prostate cancer (also called castrate-sensitive, hormone-dependent, or hormone-naive prostate cancer); N/A, not applicable; PRO, patient-reported outcome; QALY, quality adjusted life-year; RCT, randomised controlled trial

Inclusion criteria for the HRQOL review match those for the clinical effectiveness review, with the exception of the criteria shown in Table 25. The review identified 26 publications from 15 studies (all RCTs) reporting on HRQOL, patient reported outcomes (PROs) or utilities derived from disease-specific and generic PRO instruments. Studies by Jolly 2010[52] and Patrick-Miller 2016[53] were used in the ERG's critique of the company's economic model.

Table 25 Inclusion and exclusion criteria for the systematic review of HRQOL (reproduced from Table 24, Appendix D of the company submission)

Category	Inclusion criteria	Exclusion criteria	Exclusion code
Outcomes	The review will be limited to publications that report on the following outcomes: HRQL and related PROs QALYs Utilities	Publications that only report data on the following types of outcomes: Pharmacokinetics/pharmacodynamics Clinical efficacy Clinical safety Cost and resource use Epidemiological outcomes ICERs and other cost-effectiveness outcomes	Relevant outcomes unreported
Study designs	The review will be limited to publications of studies with the following designs: Prospective non-randomised controlled interventional studies Prospective longitudinal observational studies Retrospective longitudinal observational studies Cross-sectional studies RCTs	Publications of studies with the following designs: Animal studies In vitro/ex vivo studies Gene expression/protein expression studies Economic models and trial-based economic analyses	Study design not of interest
Date	2005 – 2017, inclusive	Publications published before 2005	Date
Duplicate	N/A	Publications that are duplicates of other publications in the search yield	Duplicate

Category	Inclusion criteria	Exclusion criteria	Exclusion code
Publication types	N/A	Publications of the following types: Narrative publications Non-systematic reviews Case studies Case reports Editorials	Publication type not of interest
Other criteria	Only English language articles/conference abstracts will be included	Journal articles and conference abstracts without English full-text	Non-English
Key: Abbreviations: N/A, not applicable; PRO, patient-reported outcome; QALY, quality adjusted life-year; RCT, randomised controlled trial.			

5.1.3 What studies were included in the cost effectiveness review and what were excluded? Where appropriate, provide a table of identified studies. Please identify the most important cost effectiveness studies.

In response to a clarification request by the ERG, the company provided the list of studies included in the cost-effectiveness review, which is reproduced as Table 26 below.

Table 26 Studies included in the company's systematic review of cost-effectiveness, reproduced from the company's response to ERG clarification B1

	Reference
1	Penson DF, Ramsey S, Veenstra D, Clarke L, Gandhi S, Hirsch M. The cost-effectiveness of combined androgen blockade with bicalutamide and luteinizing hormone releasing hormone agonist in men with metastatic prostate cancer. J Urol. 2005;174(2):547-52; discussion 52.
2	Ramsey S, Veenstra D, Clarke L, Gandhi S, Hirsch M, Penson D. Is combined androgen blockade with bicalutamide cost-effective compared with combined androgen blockade with flutamide? Urology. 2005;66(4):835-9.
3	Chau A, de Lemos M, Pickles T, Blood P, Kovacic L, Abadi S, et al. Use of combined androgen blockade for advanced prostate cancer in British Columbia. Journal of Oncology Pharmacy Practice. 2010;16(2):121-6.
4	Iannazzo S, Pradelli L, Carsi M, Perachino M. Cost-effectiveness analysis of LHRH agonists in the treatment of metastatic prostate cancer in Italy. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011;14(1):80-9.
5	Grabner M, Onukwugha E, Jain R, Mullins CD. Racial variation in the cost-effectiveness of chemotherapy for prostate cancer. The American journal of managed care. 2011;17(5 Spec No):e151-9.
6	Lu L, Peters J, Roome C, Stein K. Cost-effectiveness analysis of degarelix for advanced hormone-dependent prostate cancer. BJU international. 2012;109(8):1183-92.
7	Lee D, Porter J, Gladwell D, Brereton N, Nielsen SK. A cost-utility analysis of degarelix in the treatment of advanced hormone-dependent prostate cancer in the United Kingdom. Journal of medical economics. 2014;17(4):233-47.
10	Zheng HR, Wen F, Wu YF, Wheeler JRC, Li Q. Cost-effectiveness analysis of additional docetaxel for metastatic hormone-sensitive prostate cancer treated with androgen-deprivation therapy from a Chinese perspective. European journal of cancer care. 2017;26(6).

5.1.4 What does the review conclude from the data available? Does the ERG agree with the conclusions of the cost effectiveness review? If not, provide details.

A key parameter for the cost effectiveness modelling is the quality of life decrement for those in the DOC+ADT arm once they have completed their course of docetaxel: ADT (post DOC+ADT). The company derives this value from a company commissioned TTO study that compares DOC+ADT with ADT. The health state descriptors of the TTO study have been supplied at the request of the ERG. They may be biased.

Appendix H of the submission presents the details of the company systematic review of quality of life studies and associated data extraction. This is not particularly accessible and does not present the conclusions of the studies from which data have been extracted. The presentation of the results of the company systematic review of quality of life studies within the main body of the submission is insufficient for an assessment of the reasonableness of the health state descriptors of the company commissioned TTO study.

The company systematic review of quality of life studies identifies two mHPSC studies with RCT trial data for a comparison of the quality of life of DOC+ADT with ADT. One uses the EORTC-QLQ-C30 questionnaire.[52] It concludes that while DOC+ADT is associated with an initial deterioration, at 12 months there is no difference in overall quality of life between DOC+ADT and ADT. The other[53] uses the FACT-P questionnaire. It concludes that both arms resulted in some increased symptoms over time, but DOC+ADT not only provided a survival benefit but also preserved a better quality of life for mHSCP patients for longer than ADT alone. The FACT-P total score analysed with a mixed effects model estimated a net difference between the arms at baseline of -1.00 (p=0.43) in favour of ADT, with this falling further in favour of ADT to -3.09 (p=0.02) at 3 months but improving steadily thereafter to reach 2.85 (p=0.04) at 12 months in favour of DOC+ADT. This is written up in more detail in the 2018 paper by Morgans et al.[54]

The recent 2018 paper by Morgans et al[54] analyse quality of life among an RCT of DOC+ADT (n=397) compared to ADT for mHSPC (n=393). Quality of life was assessed at baseline and 3 monthly to 12 months using FACT-P, FACT-Taxane, Functional Assessment of Chronic Illness Therapy-Fatigue and the Brief Pain Inventory with the data being analysed using a mixed effect model. FACT-P completion rates were high at 90%, 86%, 83%, 78% and 77% at the five timepoints, non-completions being roughly equally split between those not given the form by staff and for unknown reasons. DOC+ADT FACT-P scores were

significantly lower at 3 months (-3.09, p=0.02) but significantly higher at 12 months compared to ADT (+2.85, p=0.04). But differences did not exceed the minimum clinically meaningful change at any time point, which was taken to be a change of 6 to 10 points. Both arms reported significantly poorer FACT-Taxane scores compered to baseline. Brief pain inventory scores were similar between the arms. The authors conclude that "*Although ADT+D was associated with statistically worse QOL at 3months, QOL was better at 12months for ADT+D patients than for ADT patients. Both arms reported a similar minimally changed QOL over time, suggesting that ADT+D is not associated with a greater long-term negative impact on QOL*".

[redacted]The company do not reference minimum clinically meaningful changes and conclude that "*Results of the ITC showed treatment with AAP+ADT was associated with notable benefits in HRQL compared to DOC+ADT. These benefits were observed from three months and sustained for at least one year after treatment*".

A crude reading of the company ITC and the results of Morgans et al[54] suggests that the 12 month FACT-P improvement from AAP+ADT compared to ADT is roughly double that of the improvement from DOC+ADT compared to ADT.

The model requires estimates for quality of life increments or decrements relative to ADT for patients in rPFS. For rPFS specific estimates of FACT-P changes there may be some confounding between both AAP+ADT and ADT and DOC+ADT and ADT in the RCT data due to more progression with ADT than with either AAP+ADT or DOC+ADT.

However, given the greater rPFS superiority for AAP+ADT over ADT compared to DOC+ADT over ADT, any such confounding might be expected to benefit AAP+ADT more than DOC+ADT. Yet, it cannot be unambiguously stated that the literature concludes that FACT-P changes for those remaining in rPFS are better among AAP+ADT patients than among ADT (post DOC+ADT) patients, or that they are better among ADT (post DOC+ADT) patients than among ADT patients.

The company have not explored the possibility of mapping from FACT-P to quality of life using the LATITUDE data as a possible means of exploring estimates based upon RCT data for AAP+ADT, DOC+ADT and ADT (post DOC+ADT) relative to ADT. It is also unclear to the ERG whether any of the three FACT-P mapping functions identified in the HERC mapping studies database[55] could help to inform this.

In the opinion of the ERG, the RCTs' quality of life data cast doubt on the company TTO study health state descriptors which assume that the quality of life among ADT (post DOC+ADT) patients is unambiguously worse than the quality of life among ADT patients. The evidence presented by the company for this unambiguous assumption also seems quite thin.

5.2 Summary and critique of manufacturer's submitted economic evaluation by the ERG

5.2.1 NICE reference case checklist

Table 27 NICE reference case checklist

Attribute	**Reference case and TA Methods guidance**	**Does the *de novo* economic evaluation match the reference case**
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	The model compares: • AAP+ADT • ADT • DOC+ADT
Patient group	As per NICE scope. "*Adults with newly diagnosed high risk metastatic hormone-naïve prostate cancer*".	In part. The data taken from LATITUDE reflects the patient population, and is analysed using multi state modelling (MSM) to derive the main transition probability matrices (TPMs) of the model. But for the company base case the outputs of the TA387[31] DES

		model for mCRPC are used as inputs. This is a poor fit due to the TA387[31] patients having a better prognosis than mHSPC patients who progress to mCRPC. The company compensates for this by applying an ad hoc hazard ratio of 2.62 to the survival probabilities derived from the TA387 model outputs.[31]
Perspective costs	NHS & Personal Social Services	Yes.
Perspective benefits	All health effects on individuals	Yes.
Form of economic evaluation	Cost-effectiveness analysis	Yes. Cost-utility.
Time horizon	Sufficient to capture differences in costs and outcomes	20 years. This is effectively a lifetime horizon.
Synthesis of evidence on outcomes	Systematic review	Yes. A systematic review and indirect treatment comparison is undertaken for mHSPC and for mCRPC.
Outcome measure	Quality adjusted life years	Yes.
Health states for QALY	Described using a standardised and validated instrument	The LATITUDE quality of life data is EQ-5D-5L.
Benefit valuation	Time-trade off or standard gamble	Time trade off.
Source of preference data for valuation of changes in HRQL	Representative sample of the public	The LATITUDE EQ-5D -5L data is cross walked to EQ-5D-3L using the van Hout et al[56] algorithm which the company describes as being recommended by the DSU. The ERG assumes this is valued using the UK social tariff, but omitted to ask this during clarification. The company has commissioned a stand-alone TTO study that estimates how much worse the quality of life for those in the

		DOC+ADT arm is compared to those in the ADT arm.
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes.
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes.
Probabilistic modelling	Probabilistic modelling	Yes. The outputs of the TA387[31] model that are used as inputs to the MSM/TA387 model of the base case are not treated probabilistically.
Sensitivity analysis		A range of univariate sensitivity analyses and scenario analyses are presented by the company. No scenario analyses limiting the duration of effect as per section 5.1.16 of the NICE methods guide are provided.

The company outline that all other companies submitting in the area have adopted a partitioned survival analysis. The company model is a quite complex Markov model. It is also unusual in having the option of applying the curves outputted by discrete event simulation model of TA387 for mCRPC as, in a sense, axiomatic inputs to the current model.

The model that is based upon the MSM analysis of the LATITUDE data augmented with clinical data from the COU-AA-302 trial will be referred to as the MSM model. The model that is based upon the MSM analysis of the LATITUDE data that also uses the output of TA387 DES model as inputs will be referred to as the MSM/TA387 model.[31] The company chooses the MSM/TA387 model for its base case.[31]

The ERG raise a number of issues with the current company model. The ERG is particularly concerned about the handling of the costs and benefits of 1st line treatment for mCRPC among patients who have progressed from their mHSPC. These are central to the cost effectiveness estimates because for AAP+ADT they provide net cost offsets to the mHSPC abiraterone drug costs.

The ERG have not attempted to address its concerns about the handling of 1st line mCRPC costs and benefits. To do so requires extensive remodelling to the extent that the major part of the model would be an ERG model rather than a company model. Moreover, it is not responsibility of the ERG to conduct such extensive remodelling.

There are some minor issues which do not much affect the current cost effectiveness estimates. These are only briefly alluded to in order to highlight the issues to the company. The issues are more simply understood through the ERG revised company model, which contains full cell referencing.

The company base case relies upon rPFS as the definition of progression. The company model also contains an option to define progression as time to subsequent therapy. The company place relatively little stress on this option. Given time constraints the ERG have not much reviewed it and has not rebuilt the model underlying it. The cost effectiveness estimates of the model that uses time to subsequent therapy as the measure of progression are more favourable for AAP+ADT than those of the company base case which uses rPFS as the measure of progression.

5.2.2 Model structure

The following covers the modelling of AAP+ADT and ADT. The modelling of DOC+ADT essentially applies the hazard ratios of the company mHSPC ITC for DOC+ADT compared to AAP+ADT to the AAP+ADT probabilities, as described in greater detail at the end of this subsection.

The company develop a *de-novo* Markov model with a weekly cycle for the 1st year and a four weekly cycle thereafter. This has three main health states:

- Progression free survival (rPFS) when patients are in mHSPC;
- Post progression survival when patients are in mCRPC; and,
- Dead.

On the basis of differences in the cumulative log hazard plots for rPFS and overall survival (OS) in the LATITUDE trial data, the company apply the LATITUDE Kaplan Meier rPFS and OS curves for the first 5 months of the model. Subsequent to the first 5 months the transition probabilities between these health states are derived from a multi-state model (MSM) statistical analysis of the post 5 months LATITUDE trial IA1 data.

The model also requires that post progression, or mCRPC, survival be split into:

- Pre 1st line treatment for mCRPC;
- On 1st line treatment for mCRPC;
- Off 1st line treatment and prior to 2nd line treatment for mCRPC;
- On 2nd line treatment for mCRPC; and,
- On 3rd line treatment for mCRPC.

Within this, 1st line treatment for mCRPC is assumed to be largely composed of active treatment, though a small proportion who are "On treatment" only receive BSC. Larger proportions only receive BSC at 2nd line, while at 3rd line virtually all patients are assumed to only receive BSC.

For the MSM model the mCRPC survival is derived from the LATITUDE MSM probabilities. The arm specific probabilities of moving from mCRPC onto 1st line treatment for mCRPC are derived from the mean treatment free intervals in the LATITUDE trial. The other probabilities that split up mCRPC survival are based upon mean times estimated from COU-AA-302 trial data.

The company argue that treatments for mCRPC during the LATITUDE trial do not reflect UK practice. As a consequence, the LATITUDE data do not reflect the relevant mCRPC survival or the probabilities splitting up PPS survival. The company model has the option to model mCRPC survival and time on mCRPC 1st line treatment using the modelled survival and discontinuation curves of the discrete event simulation that the company presented for TA387. This is the MSM/TA387 model.[31]

The TA387 model yields a mCRPC OS curve and a discontinuation curves for 1st line abiraterone for mCRPC, and a similar pair of curves for 1st line placebo or BSC for mCRPC. The current model applies arm specific proportions of patients whose 1st line mCRPC treatment is abiraterone, enzalutamide, docetaxel, cabazitaxel, radium-223 and BSC. For instance, in the DOC+ADT arm no patients receive docetaxel for their mCRPC. For each 1st line active treatment for mCRPC the mCRPC OS hazard ratio for that treatment relative to abiraterone is applied to the abiraterone mCRPC OS curve to estimate that treatment's mCRPC OS curve. The arm specific 1st line mCRPC OS curve is then calculated as a weighted average of the treatment specific and BSC mCRPC OS curves.

For the base case of the MSM/TA387 model, based upon the mCRPC ITC of the company, it is assumed that all 1st line mCRPC active treatments have the same efficacy as abiraterone. This is varied in a sensitivity analysis that applies the central estimates of the mCRPC ITC of the company, [redacted].

The company finds that applying the OS curves derived from the TA387[31] model outputs causes the MSM/TA387 model not to fit the LATITUDE OS Kaplan Meier curves. Survival is overestimated due to the COU-AAP-302 mCRPC patients having a much better prognosis than the LATITUDE mHSPC patients who have progressed to mCRPC. As a consequence, the company estimates an ad hoc 2.62 hazard ratio, or "*conversion factor*", that when applied to the modelled mCRPC OS curves derived from the TA387 model minimises the difference between the MSM/TA387 model outputs and the unweighted LATITUDE OS Kaplan Meier curves.

In essence, the MSM/TA387 model coupled with the ad hoc 2.62 hazard ratio is a complicated, non-statistical way of fitting curves to the LATITUDE OS Kaplan Meier data. The MSM model and the MSM/TA387[31] model with the 2.62 hazard ratio adjustment estimate similar OS curves during the period of the LATITUDE trial. Survival estimates only really differ between them during the extrapolation period.

While the extrapolated survival curves of the MSM model and the MSM/TA387[31] model differ during extrapolation this is not the main difference between the output of the two models. The two models mainly differ in terms of the proportions of mCRPC survival spent

on 1st line mCRPC treatment, mainly costly active treatments, and spent on 3rd line mCRPC treatment, mainly the somewhat cheaper BSC.

The written submission lacks some detail, but it appears that the MSM model estimates 1st line mCRPC treatment discontinuation from the mean times spent on 1st line mCRPC treatment during the COU-AAP-302 trial.

In the MSM/TA387 model, given the 2.62 hazard ratio adjustment of the mCRPC OS curves, it is no longer sensible to apply the TA387[31] model discontinuation curves. The company revise these discontinuation curves so that the resulting proportions of mCRPC survival spent on 1st line mCRPC treatment are the same as those implied by the unadjusted TA387 model mCRPC discontinuation and OS curves.

For both the MSM model and the MSM/TA387[31] model the probabilities of ceasing 2nd line mCRPC treatment appear to be derived from mean times during the COU-AAP-302 trial. The times spent on 3rd line mCRPC treatment seem to be residuals determined by the modelled OS curves.

The above covers the modelling of the AAP+ADT arm and the ADT arm. The company also model a DOC+ADT arm. This uses the company mHSPC ITC estimates for the hazard ratios of overall survival and progression free survival, with the company choosing to apply these to the probabilities of the AAP+ADT arm. The hazard ratios are applied to the AAP+ADT Kaplan Meier, MSM and LATITUDE derived probabilities as follows:

- rPFS to dead probability: OS hazard ratio
- PPS to dead probability: OS hazard ratio
- rPFS to PPS probability: rPFS hazard ratio
- PPS to 1st line mCRPC treatment probability: rPFS hazard ratio

The mHSPC ITC hazard ratios are not applied to any of the model inputs that are derived from the TA387[31] mCRPC model.

5.2.3 Population

The modelled population reflects that of the LATITUDE trial: mHSPC patients.

5.2.4 Interventions and comparators

For the treatment of mHSPC the company compares three arms:

- AAP+ADT
- ADT
- DOC+ADT

But the comparison is of different treatment sequences. Patients who progress from mHSPC to mCRPC receive different treatments for their mCRPC depending upon which of the three mHSPC treatment arms they have come from.

In what follows AAP+ADT, ADT and DOC+ADT will refer to the three mHSPC treatment arms. However, docetaxel for mHSPC is only received for a maximum of 6 treatment cycles of 3 weeks each. For both costs and QALYs it is necessary to distinguish between mHSPC patients who are still receiving their course of docetaxel, DOC+ADT on docetaxel patients, and mHSPC patients who have completed their course of docetaxel and so are only receiving ADT, ADT (post DOC+ADT).

5.2.5 Perspective, time horizon and discounting

A 20 year time horizon, which is effectively a lifetime horizon, is applied. The perspective and discounting is as per the NICE reference case.

5.2.6 Treatment effectiveness and extrapolation

Treatment effectiveness: mHSPC

For the first 5 months of the model, due to the LATITUDE log cumulative hazard plots varying as shown in Figures 26 and 27 of Document B of the submission, the model applies the Kaplan Meier OS and rPFS curves for AAP+ADT and ADT.

Thereafter the transition probability matrices estimated through a multi-state modelling analysis of the LATITUDE post 5 month data are applied.

The model requires that arm specific probabilities of moving from PPS pre-1st line mCRPC treatment to 1st line mCRPC treatment be derived. The company state that it was not possible to derive these within the MSM analysis as it failed to converge. Instead, the company derive these from the mean treatment free period in LATITUDE. The base case uses the mean

treatment free interval among patients who progressed during LATITUDE[a]. The company supply an additional scenario analysis at clarification that restricts this estimate to patients with data for both progression and receipt of 1st line treatment for mCRPC.

Table 28 Mean months mCRPC treatment free: LATITUDE

	AAP+ADT	ADT
rPFS patients	[redacted]	[redacted]
weekly probability	[redacted]	[redacted]
rPFS/TTST patients	[redacted]	[redacted]
weekly probability	[redacted]	[redacted]

The probabilities of moving from PPS pre-1st line mCRPC treatment to 1st line mCRPC treatment are subtracted from the MSM probabilities of remaining in PPS.

The DOC+ADT probabilities are estimated by applying the hazard ratios of the company mHSPC ITC as follows:

- OS hazard ratio of 1.09 applied to:
  - rPFS to dead probability:
  - PPS to dead probability
- rPFS hazard ratio of 1.32 applied to:
  - rPFS to PPS probability:
  - PPS to 1st line mCRPC treatment probability

This results in the following weekly transition probability matrices.

Table 29 Base case weekly TPMs: AAP+ADT

From \ To	rFPS	PPS Pre-Tx	PPS 1st line Tx	Dead
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
PPS Pre-Tx		[redacted]	[redacted]	[redacted]
PPS 1st line Tx			[redacted]	[redacted]
Dead				[redacted]

[a] The company supplied a minor correction to this at clarification with an amended model. It appears not to have applied this correction to the final model submitted upon which the cost effectiveness estimates are based. Consequently the ERG retains the original estimates. This has minimal impact on results.

Table 30 Base case weekly TPMs: AAP+ADT

From \ To	rFPS	PPS Pre-Tx	PPS 1st line Tx	Dead
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
PPS Pre-Tx		[redacted]	[redacted]	[redacted]
PPS 1st line Tx			[redacted]	[redacted]
Dead				[redacted]

Table 31 Base case weekly TPMs: DOC+ADT

From \ To	rFPS	PPS Pre-Tx	PPS 1st line Tx	Dead
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
PPS Pre-Tx		[redacted]	[redacted]	[redacted]
PPS 1st line Tx			[redacted]	[redacted]
Dead				[redacted]

The probability of dying from PPS is similar for AAP+ADT and ADT. An anomaly arises in the application of the ITC OS HR of 1.09 for DOC+ADT, this resulting in a higher probability of dying from PPS than either AAP+ADT or ADT.

When the model changes to a 4-weekly cycle the probabilities off the principal diagonal are calculated as 1-(1-p)^4, with the principal diagonal being a residual so that the rows sum to 100%.

The MSM/TA387[31] model does not apply the transition probabilities for 1st line mCRPC treatment. When reviewing the above TPMs this is better seen as an absorbing health state which is then modelled separately through the TA387 model output 2.62 hazard rate adjusted OS and discontinuations curves.

Treatment effectiveness: mCRPC

For the MSM model the LATITUDE TPMs are applied, with the probabilities of discontinuing 1st line mCRPC treatment and moving onto 2nd line mCRPC treatment apparently[b] being derived from COU-AAP-302 mean treatment times.

[b] Based upon references given in the electronic model.

For the MSM/TA387 model the LATITUDE TPMs are mostly applied with the exception of those for 1st line mCRPC treatment which can be seen as being an absorbing state. These patients are then separately modelled using the TA387 model estimated OS curves and discontinuation curves. The TA387[31] model output OS and discontinuation curves are assumed to apply to 1st line mCRPC treatment with abiraterone and 1st line mCRPC treatment with placebo or BSC. The current model applies arm specific proportions of patients whose 1st line mCRPC treatment is abiraterone, enzalutamide, docetaxel, cabazitaxel, radium-223 and BSC. The OS curves for these 1st line mCRPC treatments is estimated using their mCRPC OS hazard ratio relative to 1st line abiraterone for mCRPC, applied to the 1st line abiraterone for mCRPC OS curve. The arm specific mCRPC OS curve is then estimated as the arm specific weighted average of the active and BSC 1st line mCRPC OS curves.

The company undertake a comparison of treatments' effectiveness for mCRPC as reviewed in the clinical effectiveness section above, the estimates of which are replicated below.

Table 32 Company hazard ratios for mCRPC

	Unadjusted	Adjusted
Overall Survival HR abiraterone versus:		
Radium-223	[redacted]	[redacted]
Enzalutamide	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]
rPFS HR abiraterone versus:		
Radium-223	[redacted]	
Enzalutamide	[redacted]	

[redacted]. The company assume that the active treatments for 1st line mCRPC have identical OS curves. The company model contains the facility to apply the central estimates for the OS hazard ratios to the adjusted TA387 model estimated OS curve for abiraterone.[31]

The individual treatments' 1st line mCRPC OS curves are weighted according to the following proportions, derived from an expert panel. The model also permits the LATITUDE proportions observed at IA1 to be applied in a scenario analysis. For the LATITUDE scenario

the proportions for DOC+ADT are assumed to be those of ADT, only with the ADT docetaxel use being set to zero and these patients distributed equally between abiraterone and enzalutamide in the DOC+ADT arm.

Table 33 1st line mCRPC treatment proportions

	Base Case			LATITUDE		
	AAP+ADT	ADT	DOC+ADT	AAP+ADT	ADT	DOC+ADT
BSC	10%	5%	5%	35%	25%	25%
Enzalutamide	..	35%	39%	10%	13%	39%
AAP	..	35%	39%	3%	9%	34%
Docetaxel	60%	15%	..	51%	51%	..
Cabazitaxel	..	..	12%	1%	..	..
Radium-223	30%	10%	5%	1%	2%	2%

Applying the TA387 DES model OS curves as described above within the MSM/TA387[31] model results in OS curves that are not aligned with the LATITUDE KM OS curves. As a consequence, the company fit the MSM/TA387 model OS curves to the LATITUDE KM OS curves by estimating an ad hoc OS hazard ratio for LATITUDE mCRPC patients compared to the TA387 model output OS curves. This 2.62 hazard ratio or "*conversion factor*" is arrived at by minimising the sum of the differences between the MSM/TA387 model OS curves and the LATITUDE KM OS curves.

The LATITUDE KM OS curves, the MSM model OS curves, the unadjusted MSM/TA387 model OS curves (labelled "*No CF*") and the MSM/TA387 OS curves fitted to the LATITUDE KM OS curves using the 2.62 hazard ratio (labelled "*Base*") are as below.

Figure 4 MSM model, MSM/TA387 model and LATITUDE KM OS curves

The above shows the poorness of fit of the original MSM/TA387[31] OS curves to the LATITUDE KM OS curves. Adjusting them by the ad hoc hazard ratio of 2.62 necessarily fits them to the LATITUDE KM OS curves. However, the separation between the AAP+ADT and the ADT 2.62 hazard ratio adjusted curves is also aligned with the separation between the LATITUDE KM OS curves, which is not a necessary result of the method used to fit the curves. This could be used to argue that LATITUDE patients who progress to mCRPC have a 2.62 hazard ratio of survival compared to the modelled curves of the TA387 model,[31] and so in turn to a greater or lesser extent to the mCRPC patients of the COU-AAP-302 trial.

The above also illustrates that the MSM model and the MSM/TA387[31] model estimate near identical OS curves during the period of LATITUDE. These only really diverge during extrapolation. The MSM model OS curves lie above those of the MSM/TA387 model OS curves, but with this applies less to AAP+ADT than to ADT. Consequently, the MSM model estimates a smaller survival gain from AAP+ADT over ADT than does the MSM/TA387 model.

Adjusting the TA387 modelled mCRPC OS curves by the 2.62 hazard ratio requires that the TA387 modelled mCRPC discontinuation curves also be adjusted. The company assume that

the adjusted mCRPC discontinuation curves are the same proportions of the adjusted mCRPC OS curves as in the originally modelled unadjusted TA387 mCRPC curves.

Extrapolation

The TPMs and curves as described above are applied to the end of the 20 year time horizon, effectively a lifetime horizon.

As far as the ERG can ascertain, the sum of

- the 1st line mCRPC incident patients
- minus the sum of 1st line mCRPC patients who have discontinued
- minus the sum of 1st line mCRPC patients who have died

leaves a residual that provides an estimate of those who have received 1st line treatment but are no longer receiving it. This in turn provides an estimate of the incidence of those coming off 1st line mCRPC treatment. A portion of these incident patients are assumed to receive 2nd mCRPC line treatment which appears to be based upon mean treatment times subsequent to 1st line treatment in the COU-AAP-302 trial. The proportion of patients receiving 3rd line mCRPC treatment appears to be the residual implied by the mCRPC OS curve.

The treatment proportions for 2nd line and 3rd line mCRPC have no effect upon clinical outcomes but do determine the QALYs and costs that are applied at these stages of the model.

Table 34 2nd line mCRPC treatment proportions

	Base Case			LATITUDE		
	AAP+ADT	ADT	DOC+ADT	AAP+ADT	ADT	DOC+ADT
BSC	65%	45%	60%	84%	75%	75%
Enzalutamide	..	10%	5%	4%	8%	10%
AAP	..	10%	5%	1%	4%	7%
Docetaxel	..	10%	..	3%	5%	..
Cabazitaxel	15%	5%	5%	4%	5%	5%
Radium-223	20%	20%	25%	3%	2%	2%

Table 35 3rd line mCRPC treatment proportions

	Base Case			LATITUDE		
	AAP+ADT	ADT	DOC+ADT	AAP+ADT	ADT	DOC+ADT
BSC	90%	90%	95%	96%	91%	91%
Enzalutamide	..	..	..	1%	2%	3%
AAP	..	..	..	1%	2%	3%
Docetaxel	..	..	..	1%	1%	..
Cabazitaxel	2%	1%	1%	1%	2%	2%
Radium-223	8%	9%	4%	1%	1%	1%

5.2.7 Health related quality of life

Reports of HRQOL and utility data were sought by the company by searching MEDLINE AND EMBASE (via Embase.com), MEDLINE In-Process (vis Pubmed), NHS Economics Evaluation Database (NHS EED) and HTA Database (via Cochrane Library) in September 2015 and updated in July 2017. The searches were restricted to studies published between 2005-2017and restricted to English language publications

The search strategies are documented in full in Appendix H and are reproducible however the company conducted the MEDLINE and EMBASE searches using the EMBASE.com platform which is not accessible to the ERG.

The MEDLINE and EMBASE searches combined three search facets using the Boolean operator AND: prostate cancer; abiraterone or comparator; and HROL terms. The search strategies were considered fit for purpose, including both relevant

mHSPC quality of life values: AAP+ADT and ADT

EQ-5D-5L quality of life data was collected during LATITUDE at baseline, monthly from cycles 2-13 and every 2 months thereafter until radiographic or clinical progression of disease, at the end of study treatment, and every four months until 60 months, death, loss to follow up, withdrawal or death.

The company examined the LATITUDE EQ-5D-5L data, cross walked to EQ-5D-3L, to estimate quality of life values relationship with individual variables. These were considered for inclusion in a multivariate repeated measures mixed effect model if they had a p-value of 10% or less. The list of predictors used to derive the most appropriate utility regression equation was guided by clinical opinion, identifying the factors most likely to influence patients' HRQL, and on information from prior submissions in the mCRPC setting. This may be the reason for the exclusion of the subsequent treatment variable and the off treatment

variable from the multivariate analysis despite their p values being less than 10%. Correlation between the variables was then tested, resulting in the cycle number variable being excluded from the regression as this was found to be highly correlated with other time-dependent variables. The variables for AEs and SREs were separated out by treatment line due to possible differences between the two treatments. The univariate regression and the base case multivariate regression are as below.

Table 36 LATITUDE Regressions: univariate and base case multivariate

	Univariate (s.e.)		Multivariate (s.e.)	
Age	[redacted]	[redacted]		
Baseline EQ5D	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent Tx	[redacted]	[redacted]	[redacted]	[redacted]
Intercept			[redacted]	[redacted]
Off treatment	[redacted]	[redacted]	[redacted]	[redacted]
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
AAP+ADT Tx	[redacted]	[redacted]	[redacted]	[redacted]
SAE	[redacted]	[redacted]	[redacted]	[redacted]
SAE | AA			[redacted]	[redacted]
SAE | PBO			[redacted]	[redacted]
SRE	[redacted]	[redacted]	[redacted]	[redacted]
SRE | AA			[redacted]	[redacted]
SRE | PBO			[redacted]	[redacted]
Cycle No.	[redacted]	[redacted]	[redacted]	[redacted]

To estimate quality of life values based on the above requires that the arm specific proportions of time spent having had an SAE and having had an SRE are applied: [redacted] and [redacted] for AAP+ADT and [redacted] and [redacted] ADT respectively. Taken together these result in quality of life decrements for SAEs and SREs of [redacted] for AAP+ADT and [redacted] for ADT.

The company base case does not apply the LATITUDE quality of life decrements for SAEs and SREs. The company derive a range of estimates of quality of life decrements associated with 14 SAEs and for grouped SREs, and couple these with various durations to arrive at QALY decrements. Then they apply these to rates derived from LATITUDE for AAP+ADT and for ADT and from the literature for DOC+ADT. This results in an estimated quality of life decrement for SAEs and SREs of [redacted] for mHSPC in all three arms. This decrement is an order of magnitude less than the decrements of the LATITUDE regression. It substitutes

for them, thereby raising the quality of life values of the model above those observed during LATITUDE.

The company assume that the quality of life for DOC+ADT is as per that of AAP+ADT, but with the additional decrements outlined below.

mHSPC quality of life decrements: DOC+ADT

Due to there being no quality of life values directly attributable to DOC+ADT the company commission a quality of life study from MAPI values[c]. This concludes that, among those remaining in rPFS, on average the quality of life among those receiving docetaxel for their mHSPC and among those who have received docetaxel for their mHSPC but are now only receiving ADT is unambiguously worse than that of patients who have only ever received ADT. The worse quality of life post docetaxel use is due to depression.

Health state vignettes are developed with the aid of clinical opinion, and valued using TTO and VAS by 200 members of the general public, 88 male and 112 female, recruited through "*a panel of the general public that had expressed an interest in participating in research, members of the public responding to an advert, and snowballing/word-of-mouth*." This results in the following estimates.

Table 37 Mean values of QoL study

	VAS	(s.e.)	TTO	(s.e.)
ADT	[redacted]	[redacted]	[redacted]	[redacted]
DOC+ADT	[redacted]	[redacted]	[redacted]	[redacted]
ADT (post DOC+ADT)	[redacted]	[redacted]	[redacted]	[redacted]

The VAS and the TTO values are noticeably different, but the ratios between them are more aligned. A repeated measures GEE analysis found the cubes of the TTO estimates for the three main health states to be statistically significantly different[d].

[c] This also informs some of the adverse event quality of life values.

[d] This did not use the raw TTO values but rather used the cubes of the TTO values on grounds of skew in the data.

The company base case uses the TTO values to derive a QoL decrement for those receiving docetaxel treatment for mHSPC, DOC+ADT on treatment, of [redacted] compared to ADT and a QoL decrement for those who have finished their course of docetaxel treatment for mHSPC, ADT (post DOC+ADT), of [redacted] compared to ADT.

mCRPC quality of life values

The mCRPC quality of life values similarly ignore the LATITUDE regression decrements for SAEs and SREs and use the somewhat smaller decrements derived from the literature. These are coupled with the LATITUDE quality of life regression decrement for progression to yield the quality of life estimate for those who have progressed but are yet to receive 1st line mCRPC treatment. Those receiving mCRPC treatments have additional quality of life adjustments for treatment specific SAEs and SREs rates.

The quality of life increment of 0.02 from TA387[31] is applied to those who receive abiraterone for their mCRPC in the ADT arm and DOC+ADT arm.

The TTO quality of life decrements for those who receive docetaxel for their mCRPC are not applied in the AAP+ADT arm or the ADT arm.

The quality of life values for 2nd line mCRPC and 3rd line mCRPC are assumed to be proportionate to the values that would apply were the treatment mix being received for 1st line mCRPC. These proportions are based upon the 0.830, 0.625 and 0.500 values used in TA387 resulting in ratios of 75% and 60%.

Quality of life values: summary

The quality of life values that apply within the model are as below.

Table 38 Modelled quality of life values

	ADT+ADT	ADT	DOC+ADT
mHSPC	[redacted]	[redacted]	[redacted]
mHSPC ADT (post DOC+ADT)			[redacted]
mCRPC Pre 1st line mCRPC Tx	[redacted]	[redacted]	[redacted]
1st line mCRPC Tx	[redacted]	[redacted]	[redacted]
1st line mCRPC Off Tx	[redacted]	[redacted]	[redacted]
2nd line mCRPC Tx	[redacted]	[redacted]	[redacted]
3rd line mCRPC Tx	[redacted]	[redacted]	[redacted]

5.2.8 Resources and costs

Drug and administration costs: mHSPC

The direct drug costs are largely estimated from BNF coupled with SmPC dosing.[22] The costs for abiraterone include the commercial access arrangement, but the costs for enzalutamide, cabazitaxel and radium-223 do not include their respective patient access schemes. Dosing reflects pack size and duration, and the resulting wastage among patients who come off treatment.

The cost per docetaxel dose uses a methods of moments to calculate the distribution of LATITUDE patient BSAs and thereby the number of 20mg and 80mg docetaxel vials that would be required for the LATITUDE patient group[e]. Based upon eMIT vial costs of £3.85 and £14.74 for 80mg this results in an average cost per dose of £28.04. Using the LATITUDE patient group BSA distribution results in the same £28.04 average cost.

A compliance ratio for abiraterone for mHSPC of [redacted] is calculated from LATITUDE data, and applied to the direct drug costs.

Cycle completion rates for the six cycles of docetaxel of 96%, 93%, 91%, 89%, 85% and 84% are drawn from James et al[57] and Sweeney et al[21] and applied to the £28 docetaxel drug cost per cycle. These completion rates are not applied to the £260 chemotherapy administration cost per cycle.

[e] The cost per cabazitaxel dose for mCRPC uses the same method of moments, estimating that 0.7% of patients have a BSA of at least 4.8m^2 and so require two 60mg vials per dose. With a list price of £3,696 this results in an average cost per dose of £3,722. Using the LATITUDE patient group BSA distribution suggests marginally more, 1.1%, of patients requiring two 60mg vials and an average cost per dose of £3,736.

ADT use is assumed to be equally balanced between goserelin, leuprorelin and triptorelin, with 30% of these patients also receiving bicalutamide. The average cost per injection is assumed to be £42, with a quarter of patients incurring this cost.

Planned medical resource use (MRU): mHSPC

The submission provides limited detail of the planned MRU for treatments, though notes that it is based upon a questionnaire completed by 5 clinicians, who also subsequently attended an advisory board. The electronic model contains the following planned MRU per 4 week period for mHSPC.

Table 39 Planned mHSPC MRU: clinical advisory board

		AAP+ADT		DOC+ADT		
	Cost	<3 mth	3 mth+	≤18 Wks	18+ Wks	ADT
Oncologist visit	£101	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FBC	£3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CT scan	£123	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone scan	£292			[redacted]	[redacted]	
PSA	£7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Testosterone	£1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Liver function test	£1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Kidney function test	£1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4 weekly cost		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Annual cost		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
0.15=26 wkly, 0.22=18 wkly, 0.25=16 wkly, 0.33=12 wkly 0.44= 9 wkly, 0.67=6 wkly, 1.33=3 wkly						

The above outlines how the planned MRU for AAP+ADT lessens at 3 months, in line with the SmPC. Similarly, for DOC+ADT the planned MRU lessens after 18 weeks and completion of the docetaxel course so as to be similar to that for ADT. This is with the exception of CT scan and bone scans which are both more frequent for DOC+ADT than in the other arms and increase in frequency for DOC+ADT after 18 weeks.

Unplanned MRU, SAE and SRE costs: mHSPC

For mHSPC the company derive unplanned annual MRU frequencies from the LATITUDE trial as below. DOC+ADT is assumed to incur the same unplanned MRU as AAP+ADT.

Table 40 Unplanned medical resource use: mHSPC

	Unit cost	AAP+ADT	ADT
Radiotherapy procedure	£101	[redacted]	[redacted]
Radiotherapy preperation	£288	[redacted]	[redacted]
MRI	£180	[redacted]	[redacted]
CT scan	£120	[redacted]	[redacted]
X-ray	£171	[redacted]	[redacted]
Hospitalisation	£307	[redacted]	[redacted]
Oncologist	£173	[redacted]	[redacted]
Urologist	£103	[redacted]	[redacted]
Surgery	£12,778	[redacted]	[redacted]
Emergency room	£148	[redacted]	[redacted]
General practitioner	£38	[redacted]	[redacted]
Annual cost		£1,192	£1,513

This is augmented with the adverse event frequencies taken from the LATITUDE trial for AAP+ADT and ADT, and from Gravis et al[58] for DOC+ADT which result in additional annual costs of around £630, £580 and £1,105 respectively. The higher cost for DOC+ADT is due to 32% having neutropenia which may be reasonable to apply to those receiving docetaxel but may be less reasonable to those who have completed their course of docetaxel: ADT (post DOC+ADT) patients.

mCRPC drug and administration costs

The 1st line mCRPC compliance ratios for abiraterone and enzalutamide are assumed to be 100%[f]. This seems appropriate due to the curves that they are applied to being labelled discontinuation curves. However, for docetaxel, cabazitaxel and radium-223 the company uses treatment completion rates to estimate compliance rates of 73%, 64% and 79%. Given the discontinuation curves these are applied to, they underestimate the direct drug costs of docetaxel, cabazitaxel and radium-223 for 1st line mCRPC.

As far as the ERG can ascertain, the 1st line mCRPC treatment costs are calculated as the prevalent 1st line mCRPC on treatment population multiplied by a time invariant arm specific

[f] As reviewed later, an adjustment is applied to the costs of abiraterone for 1st line mCRPC with the intention of allowing for the [redacted], but this has little to no effect and can be ignored.

weekly treatment cost[g]. These are then qualified by whether the model cycle is during the 1st year, so 1 week long, or subsequent to this, so 4 weeks long. The treatment costs relate to those who are on treatment and incurring costs. This will not address the time dependent profiles[h] of:

- Abiraterone costs, [redacted]
- Docetaxel costs, due to a maximum of 10 cycles of 3 weeks
- R-223 costs, due to a maximum of 6 treatments separated by 4 weeks
- Cabazitaxel costs, due to a maximum of 10 cycles of 3 weeks

Planned medical resource use: mCRPC

The planned MRU for mCRPC is outlined below. The unit costs that are applied are the same as for mHSPC, these being omitted below for reasons of space.

Table 41 Planned mCRPC MRU: clinical advisory board

	AAP	DOC	ADT	ENZA	R-223	CABA
Oncologist visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FBC	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
CT scan	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
Bone scan	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
PSA	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
Testosterone	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
Liver function test	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
Kidney function test	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
4 weekly cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Annual cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Applied cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
0.15=26 wkly, 0.22=18 wkly, 0.25=16 wkly, 0.33=12 wkly 0.44= 9 wkly, 0.67=6 wkly, 1.33=3 wkly						

For reasons that are not given the company have not used the values of the clinical advisory board but have rather assumed that the planned MRU for mCRPC is equal between abiraterone and enzalutamide, and between docetaxel, R-223 and cabazitaxel. Applying the values of the clinical advisory board has minimal impact upon results.

[g] The drug cost is actually split into AAP drug costs and non-AAP drug costs with the former being qualified by the incorrect adjustment factor for the AAP cycle cap, but this can be ignored for present purposes.
[h] If as seems reasonable the same docetaxel quality of life decrements should be applied for mCRPC as for mHSPC, the time dependent profile of this will also have to be taken into account.

The values for abiraterone, ADT and docetaxel for mCRPC differ from those for mHSPC. However, while the frequencies of bone scans and CT scans increases for abiraterone, the frequencies of these for docetaxel increase by a similar amount.

The values for abiraterone are not differentiated for being prior to and subsequent to 3 months. The values for docetaxel are not differentiated for being up to and subsequent to 30 weeks, up to 10 cycles being recommended for mCRPC compared to up to 6 cycles for mHSPC.

Unplanned MRU, SAE and SRE costs and QALYs: mCRPC

For mCRPC a common annual unplanned MRU cost of £1,125 is taken from TA387, and is coupled with treatment specific SAE and SRE rates to suggest the following annual cost and quality of life effects for the mCRPC treatments.

Table 42 Unplanned MRU, SAE and SRE costs and QALYs: mCRPC

	AAP	ENZA	DOC	CABA	R-223	BSC
QoL	-0.003	-0.001	-0.004	-0.006	-0.001	0.000
Cost	£1,404	£1,286	£1,750	£2,573	£1,461	£1,125

The values applied in the AAP+ADT, ADT and DOC+ADT arms are the weighted average of these amounts.

5.2.9 Cost effectiveness results

The cost estimates of the revised company base case are as below.

Table 43 Company base case cost breakdown

	AAP+ADT	ADT	DOC+ADT
mHSPC			
Drug	[redacted]	[redacted]	[redacted]
Admin	£341	£244	£1,760
mCRPC Drug and Admin			
mCRPC: 1st line	£9,109	£16,525	£18,304
mCRPC: 2nd line	£245	£364	£309
mCRPC: 3rd line	£1,322	£1,396	£697
Other			
MRU	£20,104	£15,058	£19,533
AEs	£2,446	£1,440	£2,090
Total	[redacted]	[redacted]	[redacted]

There are large additional drug costs in the AAP+ADT arm for mHSPC, but there are also quite large cost offsets for 1st line mCRPC. The 1st line mCRPC drug and administration costs provide an offset of £7,416 for the comparison with ADT and £9,195 for the comparison with DOC+ADT. This highlights the importance of the choice of which active treatments are received for 1st line mCRPC. The choice of the lower cost DOC+ADT as the main 1st line mCRPC treatment in the ADT+AAP arm is the principal reason for the size of these cost offsets, though it has no effect upon patient outcomes in the company base case[i].

The company base case results are as follows. Note that the net amounts and ICERs are for AAP+ADT versus the comparator.

Table 44 Company base case: deterministic

	LYs	QALYs		Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	4.993	3.420		[redacted]				
ADT	3.430	2.325		[redacted]	1.563	1.095	£19,066	£17,418
DOCE	4.322	2.824		[redacted]	0.672	0.596	£10,618	£17,828

ADT is estimated to result in an undiscounted overall survival of 3.43 year, with AAP+ADT extending this by 1.56 years to 4.99 years. A patient gain of 1.09 QALYs is anticipated but

[ii] There is an insignificant effect upon quality of life due to adverse events.

costs for AAP+ADT are £19,066 higher. The cost effectiveness of AAP+ADT against ADT is estimated to be £17,418 per QALY.

DOC+ADT is estimated to result in an undiscounted overall survival of 4.32 year, with AAP+ADT extending this by 0.67 years. A patient gain of 0.60 QALYs is anticipated but costs for AAP+ADT are £10,618 higher. The cost effectiveness of AAP+ADT against AAP is estimated to be £17,828 per QALY.

The probabilistic results are in line with the deterministic results with central estimates of £17,349 per QALY for the comparison with ADT and £18,168 per QALY for the comparison with DOC+ADT[j].

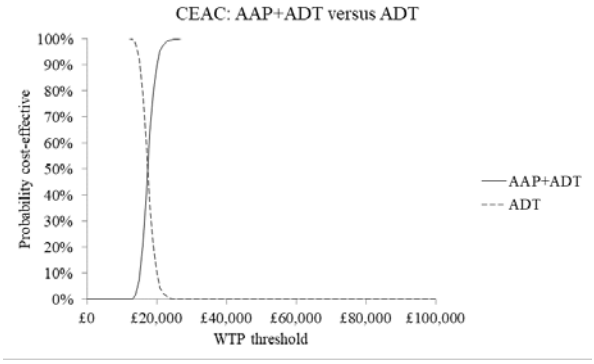


Figure 5 Company base case CEACs

5.2.10 Sensitivity analyses

Company sensitivity analyses

The company present a range of univariate sensitivity analyses, with the tornado diagrams presented as Figures 39 and 40 on page 156 of Document B of the submission and as replicated below. For data with 95% confidence intervals these were used, other parameters being varied by ±10%.

[j] The values relate to the CEACs below, which have been rerun by the ERG. The values are virtually the same as reported in Table 34 of Document B of the company submission.

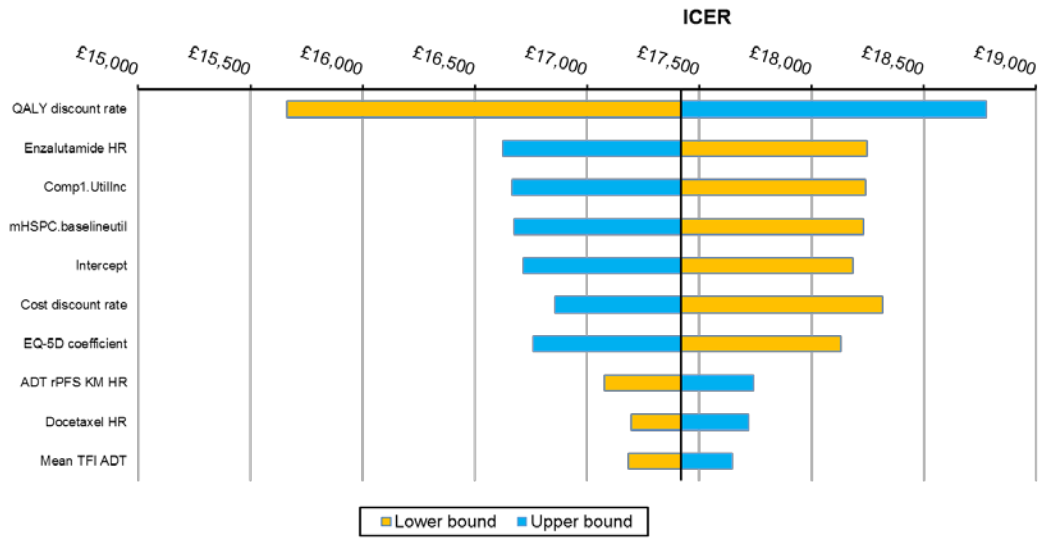


Figure 6 Company sensitivity analyses: AAP+ADT vs ADT

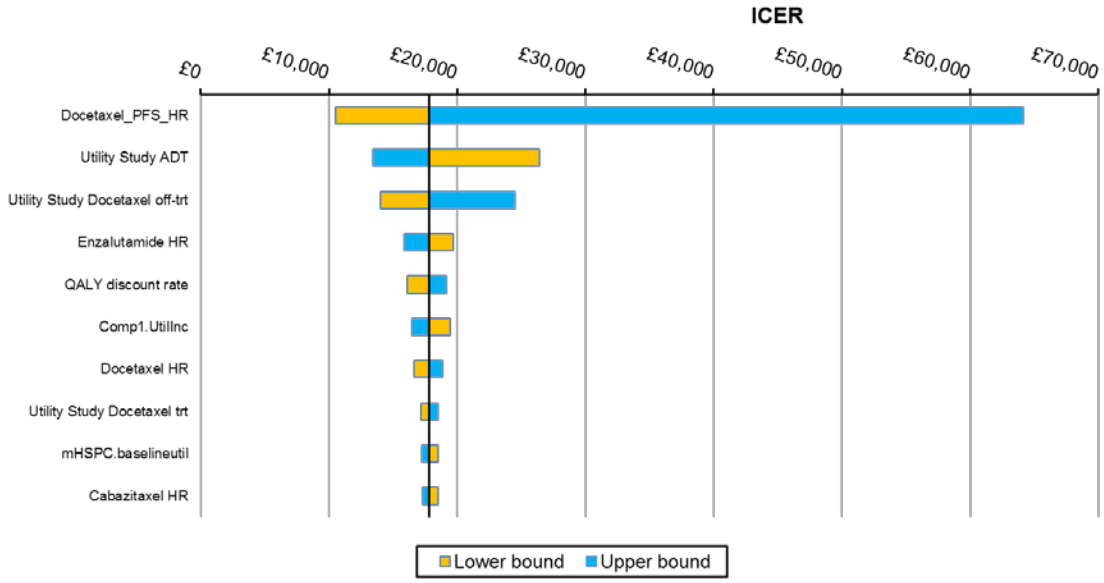


Figure 7 Company sensitivity analyses: AAP+ADT vs DOC+ADT

Company scenario analyses

The company presents a range of scenario analyses as below.

Table 45 Company scenario analyses

Model assumption	Scenario	ICER v ADT	ICER v DOC+ADT
Base Case		£17,418	£17,828
Definition of progression	TTST used as an alternative definition of progression	£14,079	£11,287
Survival and subsequent therapy source	Survival estimates and subsequent therapy market shares estimated from LATITUDE data alone	£21,504	£22,218
ITC	ITC including STAMPEDE	£17,418	£17,813
Time horizon	15 years 10 years 5 years	£17,508 £18,100 £25,856	£18,048 £19,435 £33,085
AA utility increment	Applied until death No increment applied	£16,775 £18,697	£16,656 £20,394
DOC+ADT QoL	On-treatment decrement applied only	£17,418	£20,027
AE disutilities	Using literature values alone Set to zero	£17,414 £17,361	£17,818 £17,578
mCRPC utilities	Assumed constant through mCRPC	£17,508	£17,975
AA increment (mCRPC)	AA increment from TA387 removed during mCRPC	£17,333	£17,667
Subsequent treatment ITC	Different HR are applied for each subsequent Tx based on subsequent therapy ITC	£17,129	£17,095
Vial wastage	Set to zero	£15,997	£15,077
Docetaxel cost source	MIMS price is assumed	£20,273	£16,305
AE/SRE HRQL source	Values sourced from regression	£17,510	£21,389[k]

Results show some sensitivity to:

- the time to subsequent therapy being used as the definition of progression,
- the TA387 curves being rejected in favour of just the LATITUDE MSM TPMs with this being coupled with the LATITUDE mCRPC treatment proportions,
- a time horizon of only 5 years
- applying the abiraterone quality of life increment until death
- the ADT (post DOC+ADT) TTO quality of life decrement
- vial wastage
- applying the LATITUDE QoL regression coefficients instead of a subset

Kaplan-Meier to MSM transition point

At clarification the company presented additional analyses that varied the data time point from which the MSM was performed, and so also varied the cut-off up to which the Kaplan

[kk] There is a typo in Document B of the submission, this being reported as £31,389.

Meier data was applied in the model. The ERG have updated these for the MSM/TA387 model and extended them to the MSM model as below.

Table 46 Scenario analyses around MSM start point: AAP+ADT cost effectiveness

	MSM/TA387 model		MSM model	
KM cut-off	vs ADT	vs DOC + ADT	vs ADT	vs DOC + ADT
4 months	£16,936	£17,180	£19,884	£26,001
5 months (BC)	£17,418	£17,828	£20,438	£26,909
6 months	£17,638	£18,358	£20,636	£27,619
7 months	£17,825	£19,326	£21,001	£28,545

5.2.11 Model validation and face validity check

DOC+ADT vs ADT estimates

The NICE summary of DOC+ADT compared to ADT for mHSPC states that "*In men with hormone-sensitive metastatic prostate cancer at 4 years, estimates based on a meta-analysis of the 3 RCTs (STOpCaP, n=2992)… a 9% absolute improvement in overall survival with docetaxel compared with ADT alone (49% compared with 40%, p<0.0001)… a 16% absolute improvement in time to disease progression with docetaxel compared with ADT alone (treatment failure 64% compared with 80%, p<0.0001)*".

The company base case predicts survival at 4 years of 47% for DOC+ADT compared to 34% for ADT, so a similar absolute survival for DOC+ADT but somewhat lower for ADT and hence a larger net gain of 13%.

Taking rPFS as the measure of progression the company base case predicts progression at 4 years of 75% for DOC+ADT and 87% for ADT suggesting that the model overestimates progression for both arms and particularly for DOC+ADT.

Linking the OS and rPFS together may suggest that the model overestimates the time that DOC+ADT patients spend in post progression survival. Given the importance of post progression mCRPC costs in the DOC+ADT arm for the company base case, any overestimation of the time spent in post progression in the DOC+ADT arm may of concern.

Additional ERG structural analysis

The company scenario analysis that uses the MSM model rather than the MSM/TA387 model also revises the mCRPC treatment proportions to be those of the LATITUDE trial. The company argument is that the LATITUDE data were generated by these mCRPC treatment

proportions. The results of this scenario analysis can be compared with the results of a parallel scenario analysis, but which retains the mCRPC treatment proportions of the company base case.

Table 47 Company scenario analysis: MSM model

	LYs	QALYs	Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	5.129	3.397	[redacted]				
ADT	3.597	2.302	[redacted]	1.532	1.096	£23,564	£21,504
DOCE	4.365	2.753	[redacted]	0.764	0.644	£14,312	£22,218

Table 48 ERG scenario analysis: MSM model

	LYs	QALYs	Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	5.129	3.397	[redacted]				
ADT	3.597	2.303	[redacted]	1.532	1.094	£22,356	£20,438
DOCE	4.365	2.753	[redacted]	0.764	0.644	£17,329	£26,909

The retention of the treatment proportions of the company base case has no discernible impact upon the clinical outputs of the MSM model. The net survival and net QALYs are almost unchanged. Consequently, in the opinion of the ERG when choosing which source to use for the mCRPC treatment proportions for the MSM modelling this should be driven by the need to accurately reflect the cost composition of UK mCRPC treatment patterns. By the company argument this suggests that the proportions of the company base case should be retained.

The outputs of the MSM model can also be compared with those of the MSM/TA387 model of the company base case.

Table 49 Company base case: MSM/TA387 model

	LYs	QALYs	Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	4.993	3.420	[redacted]				
ADT	3.430	2.325	[redacted]	1.563	1.095	£19,066	£17,418
DOCE	4.322	2.824	[redacted]	0.672	0.596	£10,618	£17,828

The MSM/TA387 model estimates that survival among UK mHSPC patients will be worse than that suggested by the MSM model by around 50 days for AAP+ADT patients, 60 days

for ADT patients and 16 days for DOC+ADT patients. The company MSM/TA387 model consequently suggests a larger survival gain from AAP+ADT compared to ADT but a smaller survival gain from AAP+ADT compared to DOC+ADT than the MSM model.

The net QALYs are virtually the same between the two models for the comparison of AAP+ADT with ADT. However, the net costs improve by 15% and the cost effectiveness estimate correspondingly improves by 15% when using the MSM/TA387 model.

The net QALYs for the comparison of AAP+ADT with DOC+ADT are 7.5% worse with the MSM/TA387 model than with the MSM model. This is dwarfed by the improvement in the net costs of around 40% and the cost effectiveness estimate correspondingly improves by around 35%.

As in the consideration of whether to use the LATITUDE mCRPC treatment proportions or the base case mCRPC proportions within the MSM modelling, the decision whether to use the MSM/TA387 model or the MSM model mainly affects costs. These can be further explored as below.

Table 50 MSM/TA387 and MSM model costs

	MSM/TA387 model			MSM model		
	AAP+ADT	ADT	DOC+ADT	AAP+ADT	ADT	DOC+ADT
mHSPC						
Drug	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Admin	£341	£244	£1,760	£348	£253	£1,761
mCRPC Drug and Admin						
mCRPC: 1st line	£9,109	£16,525	£18,304	£3,399	£6,513	£6,527
mCRPC: 2nd line	£245	£364	£309	£2,003	£2,673	£2,160
mCRPC: 3rd line	£1,322	£1,396	£697	£2,429	£2,709	£1,199
Other						
MRU	£20,104	£15,058	£19,533	£19,555	£14,695	£18,924
AEs	£2,446	£1,440	£2,090	£2,405	£1,440	£2,044
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The two models mainly differ in the mCRPC 1st line and 2nd line treatment costs. Both models in the above apply the same company base case mCRPC treatment proportions. Virtually all patients are assumed to receive an active 1st line mCRPC treatment, the majority

are assumed to receive an active 2nd line mCRPC treatment while the vast majority are assumed to only receive BSC at 3rd line.

The MSM/TA387 model estimates very much higher 1st line mCRPC treatment costs than the MSM model. However, the increase is less for AAP+ADT than for ADT, and is considerably less for AAP+ADT compared to DOC+ADT. The MSM/TA387 model estimates that 2nd line mCRPC treatment costs are almost negligible. These costs can be further related to the modelled undiscounted weeks spent in each health state.

Table 51 MSM/TA387 and MSM model health state durations: weeks

	MSM/TA387 model			MSM Model		
	AAP+ADT	ADT	DOC+ADT	AAP+ADT	ADT	DOC+ADT
PFS	189	106	152	189	106	152
mCRPC						
pre 1st line Tx	15	11	12	15	11	12
1st line On Tx	35	41	39	13	16	14
1st line Off Tx	1	1	1	7	8	7
2nd line	1	1	1	8	9	8
3rd line	19	19	19	35	37	34
OS Total	260	178	225	267	187	227

Within the MSM/TA387[31] model patients spend the majority of their post progression mCRPC survival in or around 1st line treatment, with around 35 to 40 weeks being spend on 1st line treatment. When these patients move on to 2nd line mCRPC treatment they are modelled as spending only around 1 week receiving it before moving into 3rd line mCRPC for around 19 weeks. The MSM/TA387 model estimates of 2nd line mCRPC duration do not seem credible.

The MSM model suggests a more evenly balanced period spent on 1st line and 2nd line mCRPC treatment, around 13 to 19 weeks for 1st line and 8 to 9 weeks for 2nd line. It also estimates a longer period at the end of survival spent at 3rd line mCRPC of around 35 weeks.

An alternative way of viewing the above is that the MSM/TA387 model estimates that patients spend around 35-40 weeks on 1st line mCRPC treatment and around 20 weeks at the end of their survival on BSC while the MSM model estimates something close to the reverse.

Fitting parameterised curves to the LATITUDE OS KM data

As noted by the company, all other companies submitting in the area have adopted a partitioned survival analysis. The ERG have not had time to explore this, and to do so would stray too far into advancing an ERG model. As a consequence, the ERG have only explored the fitting of parameterised curves to the LATITUDE OS Kaplan Meier data supplied by the company at clarification.

Up to 5 months survival in the AAP+ADT arm is that bit worse than that in the ADT arm, but is thereafter is superior to it. For an analysis assuming proportionate hazards, as per the company submission Figure 26 log-cumulative hazard plot, there is an argument that this should be restricted to the Kaplan Meier OS data subsequent to 5 months. The ERG explores both (a) using all the Kaplan Meier OS data and (b) restricting it to 5 months plus.
This suggests the following information criteria, with the Weibulls providing the best fit[l].

Table 52 ERG exploratory OS proportionate hazards analyses: information criteria

	Obs	ll(null)	ll(model)	df	AIC	BIC
All KM OS data						
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5mth+ KM OS data						
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The resulting Weibulls are as below.

[l] Analyses available to the company on request.

[redacted]

Fitting Weibulls with proportionate hazards between the arms, i.e., a common shape parameter, to all the Kaplan Meier OS data compared to just the post 5 months Kaplan Meier OS data has a reasonable effect on the extrapolated survival gains, and in particular the net gain from AAP+ADT over ADT.

Table 53 ERG's exploratory OS proportionate hazards analyses: survival estimates

	AAP+ADT	ADT	Net
Kaplan Meier "month" = 1/12 year			
All data	[redacted]	[redacted]	[redacted]
5mth +			
Kaplan Meier "month" = 4 weeks			
All data	[redacted]	[redacted]	[redacted]
5mth +			

The treatment of the Kaplan Meier data in the electronic model suggests that within the Kaplan Meier data a month relates to a calendar month[m]. This is not obviously the case, but if applied to the Weibulls estimated from the Kaplan Meier data from month 5 onwards then a partitioned survival analysis might estimate similar survival gains for AAP+ADT over ADT as both the MSM/TA387 model of the company base case and the MSM model. However, if within the Kaplan Meier data a month relates to a 4 week period the estimated survival gains

[m] If the Kaplan Meier months relate to 4 week periods this should not particularly affect the MSM/TA387 or MSM model outputs as the raw Kaplan Meier data is only applied for the first 5 months of the models, provided that the weekly MSM TPMs have correctly treated the Kaplan Meier months as 4 week periods.

for AAP+ADT over ADT fall by around 8% and are much worse than both the MSM/TA387 model and the MSM model.
The above shows how the restriction of the Weibulls to the Kaplan Meier data from month 5 onwards increases the anticipated survival gains by a reasonable margin.

The above does not address the question of whether a partitioned survival analysis would estimate a similar rPFS gain for AAP+ADT over ADT compared to both the MSM/TA387 model of the company base case and the MSM model. To do so might be for the ERG to stray too close to building an ERG model. Nevertheless, it would be relatively simple for the company to present this analysis as a confirmatory cross check of their models.

The above also does not address how a partitioned survival analysis should determine the proportions of post progression survival that are spent on 1st line, 2nd line and 3rd line mCRPC treatment. These are the main differences between the MSM/TA387 model and the MSM model.

1st line mCRPC abiraterone [redacted] *implementation*

[redacted]

[redacted]

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]		[redacted]	[redacted]
[redacted]		[redacted]	[redacted]

[redacted]

5.3 ERG cross check and critique

5.3.1 Base case results

The ERG have rebuilt the model cohort flows, QALY calculations, mHSPC costs and pre 1st line mCRPC costs using the company base case assumptions.

[redacted] The ERG have not rebuilt the 1st line, 2nd line or 3rd line mCRPC costs. In the opinion of the ERG the company model has major structural errors in the calculation of the 1st line mCRPC costs. As a consequence, there seems little point trying to rebuild them.

The ERG rebuild and the company MSM/TA387[31] model that excludes 1st line, 2nd line and 3rd line mCRPC costs result in the following undiscounted life year estimates and discounted QALY and cost estimates.

Table 55 ERG model rebuild compared to company MSM/TA387 model

	ERG Rebuild			Company MSM/TA387 model		
	AAP+ADT	**ADT**	**DOC+ADT**	**AAP+ADT**	**ADT**	**DOC+ADT**
Total						
LY	5.062	3.549	4.405	4.993	3.430	4.322
QALYs	3.455	2.379	2.863	3.420	2.325	2.824
Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Net						
LY		1.513	0.657		1.563	0.672
QALYs		1.077	0.592		1.095	0.596
Costs		£27,185	£19,195		£26,903	£19,136

The total undiscounted life years of the ERG rebuild are slightly higher than those of the company MSM/TA387 model, but the models' estimates are within 2-3% of one another. Total QALY estimates are similarly close between the ERG rebuild and the company MSM/TA387 model. The total cost estimates are little different between the models. The estimates of net amounts differ less between the models than the estimates of the total amounts.

The correspondence between the ERG rebuild and the company MSM/TA387 model is good. Nevertheless, these values cannot be taken forward to cost effectiveness estimates because they do not include 1st line, 2nd line or 3rd line mCRPC costs. These are major drivers of the cost effectiveness estimates.

5.3.2 Data inputs: correspondence between written submission and sources cited

Mapping the EQ-5D-5L to the EQ-5D-3L

The company state that the DSU recommends the van Hout mapping algorithm.[56] The DSU report on mapping the EQ-5D-5L to the EQ-5D-3L[59] states that "*The DSU and van Hout approaches… do not perform substantially differently from each other ... The DSU approach slightly outperforms van Hout in terms of predicting the category of response. The van Hout method is marginally better for some measures of summary fit to utility scores. However, we outline how these summary measures mask differences between the approaches in different parts of the health distribution. There are concerns about the validity of the pairwise deletion method employed by van Hout et al and how this distorts fit measures*." This could be read as the DSU preferring the DSU mapping method over the van Hout method. The ERG cannot

comment upon the impact that using the DSU method would have upon results. The company do not state whether this method was explored.

1st line mHSPC SAEs and SREs: DOC+ADT

The 32% for grade 3/4 hypertension with DOC+ADT corresponds with that reported in Gravis et al 2013[58] it also being necessary to note that the reported rate for ADT was 0%. The ERG has not been able to source SRE rates for DOC+ADT or for ADT from Gravis et al 2013[58] As a consequence it may not be reasonable to assume a relative risk of 91% for DOC+ADT compared to ADT, when the relative risk from LATITUDE for AAP+ADT compared to ADT is 79%.

Chemotherapy administration cost

The ERG have not been able to source the £260 chemotherapy administration cost that is applied for docetaxel administrations. The 2015-16 reference costs for outpatient administration suggest first administration costs of £265 for more complex parenteral chemotherapy and £304 for complex chemotherapy including prolonged infusion, and £212 for subsequent cycles. Applying these would have little impact upon the cost effectiveness estimates.

ADT administration cost

The ERG have been unable to source the average cost of £10.85 for ADT administrations. This has been calculated as £42*(15.5/60).

5.3.3 Data inputs: correspondence between written submission and electronic model

The ERG have not identified any important discrepancies between the written submission and the electronic model.

5.3.4 ERG commentary on model structure, assumptions and data inputs

Modelling of mCRPC

Most of the 1st line mCRPC treatments are time limited. It appears that both the MSM model and the MSM/TA387 model assume that they are not. The costs of mCRPC treatments are applied indefinitely while on 1st line mCRPC treatment. A similar issue applies to any time dependent quality of life values.

Given the centrality of the mCRPC costs to the cost effectiveness estimates, if these are incorrect it is doubtful whether any of the modelling results of the company or the ERG are reliable.

In the opinion of the ERG the time profile of 1st line mCRPC treatments' costs and quality of life values should be modelled. A possible method might be to calculate discontinuation curve adjusted present values for the mCRPC active treatments and BSC. Arm specific weighted average present values could then be calculated and applied to each cycles' incident 1st line mCRPC patients on treatment. These present values might need to be cycle specific to avoid projecting costs and benefits beyond the time horizon. The ERG have not attempted to implement this because:

- The extent of model revision would result in it being in large part an ERG model, and
- Committee has previously rejected similarly extensive model revisions by the ERG.

Choice of MSM model, MSM/TA387 model or partitioned survival analysis model

The company argue that the MSM/TA387[31] model is appropriate because:

- The LATITUDE OS data are of limited maturity.
- More patients in the ADT arm than the AAP+ADT arm of LATITUDE received subsequent mCRPC treatment.
- Some mCRPC treatments in LATITUDE are not available in the UK, and the proportions of subsequent treatments are not aligned with market shares and modelling mCRPC survival as in the MSM/TA387 model permits this to be addressed.
- A Markov model is appropriate due to the discrete event simulation of TA387 being poorly received by the Committee.

The MSM model and the MSM/TA387 model estimate very similar net survival estimates and net QALY estimates for AAP+ADT compared to ADT.

The MSM model and the MSM/TA387[31] model estimate some differences in net survival and net QALYs for AAP+ADT compared to ADT. But these are dwarfed by the differences in the net cost estimates. These in turn are driven by the proportions of mCRPC time modelled as being spent on 1st line, 2nd line and 3rd line mCRPC treatments. The choice between the MSM

model and the MSM/TA387 model largely boils down to a choice between these modelled durations. In the opinion of the ERG the MSM/TA387 model estimates for 2nd line mCRPC treatment durations lack credibility.

Applying the unadjusted TA387 OS curves results in the MSM/TA387 model fitting the LATITUDE OS Kaplan Meier curves not very well at all. The company adjust the TA387 modelled OS curves using an *ad hoc* hazard ratio of 2.62 to fit the MSM/TA387 model OS curves to the LATITUDE OS Kaplan Meier curves. This calls into question the relevance of the TA387 model outputs to mHSPC patients who progress to mCRPC.

The use of the TA387[31] model outputs as axiomatic inputs to the MSM/TA387 model may raise procedural issues. As the company note, it chose a Markov model due to the discrete event simulation of TA387 being poorly received by the Committee. The ERG also cannot be expected to have, and has not, cross checked, rebuilt, stress tested or indeed done anything very much with the TA387 model. That TA387 approved abiraterone for use for mCRPC does not imply that the Committee viewed the TA387 model outputs as the most likely central estimates that would apply in practice.

The application of the *ad hoc* hazard ratio of 2.62 to the TA387[31] modelled OS curves is essentially a laborious and non-statistical means of fitting curves to the LATITUDE OS Kaplan Meier data. If this is the intention it would be simpler to fit parameterised curves to the LATITUDE OS Kaplan Meier data. This would benefit from well-established formal statistical methods and would permit time varying probabilities to be explored. An exploratory ERG's analysis of the LATITUDE OS Kaplan Meier data suggest Weibulls are a better fit than exponentials.

Kaplan Meier cut-off

The LATITUDE Kaplan Meier curves are applied for the first 5 months of the model. The probabilities of the MSM analysis are estimated from the LATITUDE post 5 months data, and are applied in the model from 5 months.

This is a choice based upon the company examination of the log cumulative hazard plots for OS and rPFS. Viewed in isolation the log cumulative hazard plot for OS might suggest a later cut-off. Later cut-offs worsen the cost effectiveness estimates.

mCRPC treatment sequencing

There are some uncertainties around treatment sequencing, whether patients are currently only permitted one "*novel agent*" for treatment of metastatic prostate cancer and whether approval of a novel agent for mHSPC by NICE might in time increase the number of novel agents mHSPC patients will be able to receive for their metastatic prostate cancer.

The ERG accept the company argument that patients who are receiving a course of docetaxel treatment have a lower quality of life than if they were to receive a novel agent. If patients who have received abiraterone for their mHSPC can exercise choice over their treatment for mCRPC it seems likely that many if not most will prefer enzalutamide over docetaxel.

The effect of 1st, 2nd and 3rd line treatments for mCRPC on the MSM/TA387 model are not obvious from the headline results. To better understand the working of the model these can be simplified through the following 6 scenario analyses:

1. 1st, 2nd and 3rd line all receive BSC.
2. 1st line all receive enzalutamide, 2nd and 3rd line all receive BSC.
3. 1st line all receive enzalutamide, 2nd line all receive cabazitaxel, 3rd line all receive R-223.
4. 1st line AAP+ADT patients receive docetaxel while ADT and DOC+ADT patients receive enzalutamide, 2nd and 3rd line all receive BSC.
5. 1st line AAP+ADT patients receive docetaxel while ADT and DOC+ADT patients receive enzalutamide, 2nd line all receive cabazitaxel, 3rd line all receive R-223.
6. 1st line all receive enzalutamide, 2nd line AAP+ADT and ADT patients receive docetaxel while DOC+ADT patients receive cabazitaxel, 3rd line AAP+ADT and ADT patients receive cabazitaxel while DOC+ADT patients receive R-223.

These scenario analyses result in the following for the MSM/TA387 model.

Table 56 mCRPC treatment sequencing scenario analyses

	AAP+ADT vs ADT			AAP+ADT vs DOC+ADT		
	Δ QALYs	**Δ Costs**	**ICER**	**Δ QALYs**	**Δ Costs**	**ICER**
Base	1.095	£19,066	£17,418	0.596	£10,618	£17,828
SA01	1.126	£26,750	£23,752	0.619	£19,060	£30,788
SA02	1.106	£22,233	£20,095	0.608	£16,713	£27,488
SA03	1.106	£20,934	£18,920	0.608	£15,594	£25,646
SA04	1.104	£3,076	£2,785	0.606	-£2,444	Dominant
SA05	1.105	£1,776	£1,608	0.606	-£3,564	Dominant
SA06	1.107	£21,195	£19,155	0.607	£14,953	£24,625

Note that the above does not take into account the competitor PASs, and in particular the enzalutamide PAS. The 4th and 5th scenario analyses should consequently not be taken too literally, but rather as an indication of how the mCRPC treatment sequencing affects the model outputs. SA06 is also sensitive to competitor PASs due to only the 3rd line treatment for AAP+ADT and ADT being subject to a PAS while both 2nd and 3rd line treatment for DOC+ADT are subject to a PAS. The equivalent of the above table inclusive of the competitor PASs is presented in the cPAS appendix.

The net QALYs are only really affected by differentiation of the 1st line mCRPC treatment; SA02 and SA03 have the same net QALYs and those of SA04 and SA05 are little different. This is driven by changing the proportions who receive BSC rather than an active treatment due to all active 1st line mCRPC treatments being assumed to have the same efficacy.

The net QALYs increase slightly in all the scenario analyses compared to the base case. This appears to be due to the company base case assuming that in the ADT+AAP arm 90% of mCRPC patients receive an active 1st line mCRPC treatment while in the ADT and DOC+ADT arms this is 95%.

If the treatment sequences for mCRPC are similar between the arms the net costs increase and the cost effectiveness of AAP+ADT worsens markedly. This would correspond to the situation where patient choice leads to patients preferring the newer agents rather than docetaxel for their mCRPC regardless of their previous treatment for mHSPC.

The 6th scenario analysis suggests that differentiation of 2nd and 3rd line treatments for mCRPC is of secondary importance compared to differentiation of 1st line treatments for mCRPC.

MSM TPMs and application of DOC+ADT hazard ratios

The company choose to apply the DOC+ADT versus AAP+ADT hazard ratios to the AAP+ADT Kaplan Meier curves and MSM TPMs. It could have chosen to apply the DOC+ADT versus ADT hazard ratios to the ADT Kaplan Meier curves and MSM TPMs.

It appears that the MSM TPMs for AAP+ADT are estimated separately from the MSM TPMs for ADT. This is akin to parameterised curves not imposing proportionate hazards between the arms, but curves being estimated separately for each arm. In these situations applying the hazard ratios of an ITC to the curves of one of the arms of the trial will not necessarily result in the same or even similar results as applying the implied hazard ratios to the curves of the other arm of the trial.

The ERG have already highlighted that applying the DOC+ADT versus AAP+ADT hazard ratios to the AAP+ADT MSM TPMs results in the anomaly of DOC+ADT patients having a higher probability of dying once they have progressed to mCRPC than AAP+ADT patients and ADT patients.

The company could equally well have chosen to apply the hazard ratios for DOC+ADT versus ADT to the ADT Kaplan Meier curves and MSM TPMs. A crude application of the central estimates of the hazard ratios of Table 18 of Document B of the submission suggests hazard ratios of 0.67 (0.62/0.92) for OS and 0.62 (0.47/0.76) for rPFS for DOC+ADT compared to ADT. Applying these to the ADT MSM TPM results in the following TPM for DOC+ADT.

Table 57 Scenario analysis TPMs: DOC+ADT

From \ To	rFPS	PPS Pre-Tx	PPS 1st line Tx	Dead
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
PPS Pre-Tx		[redacted]	[redacted]	[redacted]
PPS 1st line Tx			[redacted]	[redacted]
Dead				[redacted]

Most of the values in the above are in line with intuition when compared with the TPMs of AAP+ADT and ADT. However, the probability of dying among those who have progressed is anomalous and is now lower than that of both AAP+ADT and ADT. It can be argued that this anomaly is worse than that of the DOC+ADT TPM of the company base case.

Application of the above TPM considerably worsens the deterministic MSM/TA387 model cost effectiveness estimate for AAP+ADT compared to DOC+ADT from £17,828 per QALY to £25,530 per QALY. The ERG implementation of sampling of this within the probabilistic modelling may be formally incorrect and may not properly take into account confidence intervals and correlations. Nonetheless, this results in a smaller change in the central probabilistic estimate, it only worsening from £18,168 per QALY to £20,867 per QALY. The non-linearity of the model may relate to the DOC+ADT versus ADT OS hazard ratio being somewhat further from unity than the DOC+ADT versus AAP+ADT hazard ratio.

The above does not argue that the company choice is incorrect. It only highlights that it is a choice which has not been justified, another choice could equally well have been made and that the most reasonable estimate may lie somewhere between the two.

MSM/TA387 model: Differentiation of 1st line mCRPC treatment effects

As already highlighted, the company comparison of 1st line mCRPC treatments' effectiveness estimates an OS hazard ratio central estimate which [redacted]

The company MSM/TA387 model structure is largely justified by the company on the basis of the need to properly model the effects of extending rPFS upon OS; i.e., the LATITUDE data for rPFS are reliable but thereafter the modelling needs to depart from the LATITUDE data. [redacted].

The only means of approximating this within the MSM/TA387 model is to differentiate 1st line mCRPC treatments by the company central estimates of the OS hazard ratios. The ERG will apply this as a sensitivity analysis.

The NICEimpact Cancer publication[60] provides prescription data for enzalutamide and abiraterone for mCRPC which suggest a strongly rising market share for enzalutamide and a falling market share for abiraterone, with a prescribing ratio of around 2:1 in April 2017. It also notes that "*Enzalutamide is similar to abiraterone, but it is less likely to cause liver toxicity and may be more convenient to take for some people*". In the light of this the ERG sensitivity analyses that differentiate 1st line mCRPC treatment effectiveness will apply the treatment proportions of the ERG base case and a second set of proportions that sets the proportion of abiraterone to zero and adds this to the proportion for enzalutamide in the ADT and DOC+ADT arms.

MSM model: What proportions of mCRPC treatments to apply?
The company scenario analysis which uses the MSM model also applies the LATITUDE mCRPC treatment proportions. The implicit company argument appears to be that it is these treatment proportions that gave rise to the LATITUDE clinical data, so these should be applied in the mainly LATITUDE based MSM model.

The company argument might be reasonable if the MSM model survival estimates and QALY estimates are sensitive to the mCRPC treatment proportions. They are not. In the light of this the ERG consider it more important to accurately estimate the costs of mCRPC treatment. According to the company, this is best achieved using the estimates of the company's clinical Advisory Board.

The ERG are unclear whether the proportions receiving BSC for 1st line mCRPC should be differentiated between the arms. For the ERG revised base case the ERG make minor amendments to the company clinical advisory board estimates as below.

Table 58 ERG revised 1st line mCRPC treatment proportions

	AAP+ADT	ADT	DOC+ADT
1st line mCRPC			
BSC	5%	5%	5%
Enzalutamide	0%	35%	39%
AAP	0%	35%	39%
Docetaxel	65%	15%	0%
Cabazitaxel	0%	0%	12%
Radium 223	30%	10%	5%
2nd line mCRPC			
BSC	60%	45%	60%
Enzalutamide	0%	10%	5%
AAP	0%	10%	5%
Docetaxel	0%	10%	0%
Cabazitaxel	15%	5%	5%
Radium 223	25%	20%	25%
3rd line mCRPC			
BSC	95%	90%	95%
Enzalutamide	0%	0%	0%
AAP	0%	0%	0%
Docetaxel	0%	0%	0%
Cabazitaxel	2%	1%	1%
Radium 223	3%	9%	4%

The ERG provide a scenario analysis that applies the company Advisory Board estimates.

LATITUDE QoL Regression

The univariate model estimates *Off Treatment* and *Subsequent Treatment* to be significant at not just the 10% level but at the 5% level. Despite this they are excluded from the multivariate analysis. The univariate coefficients for both of these are somewhat larger than those of the other variables, with the exception of the *Baseline EQ5D* and *SREs*. The ERG asked the company to supply the internal reports that underlay the estimates reported in the submission, but none were forthcoming.

In the light of the above, the ERG cannot comment further upon why *Off Treatment* and *Subsequent Treatment* were excluded from the multivariate analysis, how justified it was to

exclude them and what the impact of including them would have been. But the following can be noted.

- The *Off Treatment* health state is intrinsic to the model structure. When calculating the quality of life value for those on AAP+ADT the increment for receiving AAP+ADT of [redacted] is qualified by the [redacted] of time prior to cessation of therapy that patients actually receive AAP+ADT. Inclusion of the *Off Treatment* variable in the regression equation might reduce the quality of life during this period in the AAP+ADT arm.
- The amount of time spent *Off Treatment* prior to treatment for mCRPC in the MSM/TA387 model[n] is 3.4 months for AAP+ADT, 2.5 months for ADT and 2.8 months for DOC+ADT. Inclusion of the *Off Treatment* variable in the regression equation might reduce the quality of life during this period in the AAP+ADT arm more than that in the other arms.
- The distinction between those with rPFS who are receiving subsequent treatment and who are not is also inherent to the model. The quality of life differences between these is modelled as being minimal and only due to the adverse events associated with the various treatments. Whether the inclusion of the *Subsequent Treatment* variable within the analysis would increase this difference is unclear.

LATITUDE QoL Regression: Differentiating SAE and SRE effects by arm

The ERG requested additional analyses that variously pooled the coefficients for SAEs and for SREs between the arms, and asked what statistical justification there is for separating them by arm. It supplies the following models for the rPFS QoL analysis.

[n] Taken to be the sum of the elements of Column J of the markov worksheets, conditioned by 0.23 months to cycle 52 and then 0.92 months.

Table 59 LATITUDE QoL regressions: pooling of coefficients

	Base case (s.e.)	Pooled SAE (s.e.)	Pooled SRE (s.e.)	Both pooled (s.e.)
Baseline EQ5D	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]
rPFS	[redacted]	[redacted]	[redacted]	[redacted]
AAP+ADT Tx	[redacted]	[redacted]	[redacted]	[redacted]
SAE		[redacted]		[redacted]
SAE | AA	[redacted]		[redacted]	
SAE | PBO	[redacted]		[redacted]	
SRE			[redacted]	[redacted]
SRE | AA	[redacted]	[redacted]		
SRE | PBO	[redacted]	[redacted]		
-2 Res LL	[redacted]	[redacted]	[redacted]	[redacted]
AIC	[redacted]	[redacted]	[redacted]	[redacted]
AICC	[redacted]	[redacted]	[redacted]	[redacted]
BIC	[redacted]	[redacted]	[redacted]	[redacted]

All coefficients are significant at the 1% level.

At clarification the ERG asked to what extent there was statistical evidence that the SAE and SRE coefficients differed by arm. The company note that:

- *"LATITUDE evidence suggests that the impact of having experienced an AE or SRE was different in the AAP + ADT arm than in the ADT alone arm. The utility regression analysis highlighted some difference, with the coefficient for AE being -[redacted] for AAP + ADT and [redacted] for ADT alone, and the coefficient for SRE being -[redacted] for AAP + ADT and [redacted] for ADT alone."*
- *"Each of the variables included in the utility regression model 1.0, which estimates treatment-specific AE and SRE coefficients, were found to be statistically significant … The p-values for the AE and SRE coefficients separated by treatment arm are all well below 0.01."*

The above arguments examine the arm specific coefficients in isolation and do not address whether the coefficients are statistically different between the arms. Pooling the SAE coefficients quite noticeably improves the information criteria, though further pooling the SRE coefficients provides little additional gain. The company maintain that the cost effectiveness estimates are not sensitive to which model is chosen, largely because the pooling of SAE and SRE coefficients is balanced by an increase in the AAP+ADT treatment effect coefficient.

The ERG will apply the coefficients of the model that pools the SAE coefficients due to the improvements in the information criteria, slightly improving the cost effectiveness estimates for AAP+ADT. However, as the company notes the choice of model from those available has relatively little impact upon the cost effectiveness estimates due to the AAP+ADT treatment effect coefficient increasing.

Selective application of the LATITUDE QoL Regression

As already noted the company chooses not to apply the QoL decrements for SAEs and SREs that are implied by the LATITUDE QoL regression. It estimates decrements from the literature that are an order of magnitude smaller than those implied by the LATITUDE QoL regression. This causes the quality of life estimates used in the model to be higher than those observed during LATITUDE. This biases the model in favour of AAP+ADT.

The ERG can think of no reason for adopting this approach for the comparison of AAP+ADT with ADT. There might be an argument for qualifying the LATITUDE QoL regression decrements for SAEs and SREs for AAP+ADT before applying them to DOC+ADT if the literature estimates suggested wildly differing values. However, the company estimates based upon the literature are minimally different, a decrement of [redacted] for AAP+ADT compared to [redacted] for DOC+ADT.

The ERG will apply the LATITUDE regression in full in its revised base case, this also applying the minor qualification to the DOC+ADT decrement implied by the literature based estimates of the company for AAP+ADT and DOC+ADT.

DOC+ADT and ADT (post DOC+ADT) QoL compared to ADT QoL: mHSPC

The company commission a QoL study from MAPI values. This develops three main health states based upon literature review and the input of 4 patients, 3 expert clinicians and 2 nurses. An additional 6 health states are developed by adding adverse events to one of the main health states.

The study notes that

“[redacted]

[redacted] [ERG emphasis]".

The three main health states are ADT, DOC+ADT while on docetaxel treatment and ADT after having completed a course of docetaxel. The full health descriptions are presented in Appendix 1, but for reasons of space and clarity only the elements that differed are presented below.

Table 60 Quality of life study health state descriptors

ADT	DOC+ADT	ADT (post DOC+ADT)
[redacted]	[redacted]	[redacted]

These health states were further reviewed by 5 clinicians and 1 nurse who had not been previously involved in the study, a key question being whether patients were more depressed following a course of docetaxel. The 6 experts were equally split, with 3 reporting that some patients were more depressed following docetaxel treatment. It is unclear whether they thought that these patients were more depressed due to the docetaxel treatment or simply due to having had a longer duration of disease.

In the opinion of the ERG the heath state descriptors for DOC+ADT and ADT (post DOC+ADT) are unambiguously worse than the health state descriptor for ADT. It is inevitable that when valued by members of the public they will result in quality of life decrements for DOC+ADT and ADT (post DOC+ADT) compared to ADT

The first criticism of the QoL study is that it did not investigate any AAP+ADT health state. This would provide an estimate of the QoL detriment for ADT compared to AAP+ADT, and so some cross check about the alignment and reliability of the study estimates against estimates based upon trial data and real patients' experiences.

The health state descriptors may be biased against the DOC+ADT and ADT (post DOC+ADT) health states.

- In the opinion of the ERG "[redacted]" seems likely to be viewed as having a better prognosis than "[redacted]"[o]. To the ERG the former suggests the possibility of treatment while the latter suggests something immutable and unchanging. There is no justification for the difference in wording between ADT and DOC+ADT. It also seems questionable whether this wording should be differentiated between AAP+ADT and ADT (post DOC+ADT) given that the quality of life decrements are applied for the duration of mHSPC.
- For the ADT (post DOC+ADT) health state it is not obvious why "[redacted]" needs to be included in the health state description. To the ERG it seems questionable whether members of the general public can sensibly infer what effect past treatment as specified in the health state descriptor will have on their quality of life, but its inclusion seems likely to push responses by members of the public in only one direction. The anticipated effects of this would seem to be covered by the subsequent depression related wording, which does form what can reasonably be described as part of someone's health state.
- For the DOC+ADT health state "[redacted]" seems to overstate the

[o] [redacted]

restriction on daily activities given that docetaxel administration is only once every three weeks. The patient who reported on this also only restricted his social activities during the week he received treatment.

- There seems to be considerable uncertainty about whether there is a difference in depression for ADT (post DOC+ADT) compared to DOC+ADT, and indeed compared to ADT. This uncertainty is not reflected in the wording of the health state for ADT (post DOC+ADT) for which depression is unambiguously "[redacted]" rather than "[redacted]" for ADT and DOC+ADT.
- It seems peculiar to assume that depression among patients worsens when they complete their course of docetaxel.
- It would have been simple to include an indication of median future survival within the health state descriptors. The better prognosis for ADT (post DOC+ADT) patients than for ADT patients at the same time point is not reflected in the health state descriptors. The better prognosis for ADT (post DOC+ADT) patients than for ADT patients at the same time point may result in them being less depressed.

Given the uncertainty around the likelihood of increased depression for the ADT (post DOC+ADT) compared to ADT, it might be better to explore this as an adverse event rather than as an inseparable aspect of ADT (post DOC+ADT) health state.

In the light of the above and the FACT-P values reported in section 5.1.4 above, the ERG revised base case will set the quality of life decrement for ADT (post DOC+ADT) compared to ADT to zero. The ERG will apply the [redacted] TTO decrement within sensitivity analyses. In the light of the FACT-P results of section 5.1.4 above, the ERG will also apply half the LATITUDE quality of life regression increment for AAP+ADT to ADT (post DOC+ADT) within a sensitivity analysis.

DOC+ADT and ADT (post DOC+ADT) QoL compared to ADT QoL: mCRPC

The company does not apply the quality of life decrements for DOC+ADT and ADT (post DOC+ADT) QoL compared to ADT for those treated with docetaxel for mCRPC in either the AAP+ADT arm or the ADT arm. This biases the model in favour of AAP+ADT when compared to DOC+ADT, to a lesser extent in favour of AAP+ADT when compared to ADT.

Given the ERG's concerns around the modelling of time dependent mCRPC costs the ERG have not attempted to address this in their revised base case. The same issue applies to time dependent mCRPC quality of life values.

AAP+ADT abiraterone treatment compliance estimate: mHSPC

The company calculate that patients in the AAP+ADT arm receive treatment with AAP+ADT for [redacted] of the time they spend in rPFS. While the percentage affects both costs and QALYs in the AAP+ADT arm, it reduces costs more than QALYs. Not applying the percentage reduction worsens the costs effectiveness estimate for AAP+ADT compared to ADT[p] from £17,418 to £20,038 per QALY, and for AAP+ADT compared to DOC+ADT from £17,828 to £22,593 per QALY. The [redacted] estimate is essentially based upon the differences in the areas underneath the rPFS and TTD curves as outlined below.

[redacted]

[redacted]

The areas under the curves are around [redacted] months for the rPFS KM curve and [redacted] months for TTD KM curve, which results in a ratio of [redacted] time on treatment compared to time in rPFS.

[p] This has also set the corresponding proportion for ADT to 100%.

The above figure raises concerns about the [redacted] compliance ratio.

- Some of the separation between the curves appears to be protocol driven, due to the rPFS curve being flat for 4 months and followed by a step at the assessment point.
- Over the 40 months the sum of events and censoring events was the same for both curves at [redacted]. But the balance between events and censoring events was considerably lower in the rPFS curve at [redacted] compared to [redacted] for the TTD curve. The definition of censoring events for the rPFS curve was broader than for the TTD curve. If the ERG understand the company clarification response correctly, for the TTD curve only remaining on treatment at IA1 counted as censoring with all other events being TTD events. Consequently, lost to follow up is treated as censoring for the rPFS curve, which is unaffected by it, whereas it is treated as an event in the TTD curve, causing it to fall. Like may not be being compared with like.
- The [redacted] ratio does not take into account the numbers at risk. At baseline the ratio is near 100% and almost all patients remain at risk. By 40 months there are virtually none at risk and the ratio between the curves has dropped to around [redacted]. The 100% and [redacted] are given equal weight.

If the company have confidence in the curves there is a clear downward trend in the ratio as time passes. The company should extrapolate from this steeply downward sloping curve. It has not done so. This suggests that the company do not find the end of curve ratios credible.

There does not seem to be an agreed method for handling this, what censoring should be informative and what uninformative in the TTD curve and how any estimate should be qualified by the numbers at risk. But there may be no need to address these issues. The CSR contains data on treatment compliance in the safety population as outlined below, with this being described as "*Percent of doses (tablets) taken out of the protocol-specified dose*".

Table 61 CSR compliance data for AAP/placebo

Compliance range		AAP+ADT		ADT	
	75%	[redacted]	[redacted]	[redacted]	[redacted]
>75%	80%	[redacted]	[redacted]	[redacted]	[redacted]
>80%	85%	[redacted]	[redacted]	[redacted]	[redacted]
>85%	90%	[redacted]	[redacted]	[redacted]	[redacted]
>90%	95%	[redacted]	[redacted]	[redacted]	[redacted]
>95%	100%	[redacted]	[redacted]	[redacted]	[redacted]

The ERG may have misinterpreted this data, which could explain the more convoluted approach of the company.

Given the clear rightward skew in the compliance data, it seems reasonable to assume a similar skew within each of the ranges. There is the problem of the 1st compliance range of up to 75%, and the ERG have little option other than to treat this as 75%. Taking the upper limit of the other ranges results in mean compliances of [redacted] for AAP+ADT and [redacted] for ADT, while the midpoints result in mean compliances of [redacted] for AAP+ADT and [redacted] for ADT.

The rightward skew may argue for the upper limit estimates to be used. But given the difficulty around the 1st compliance range the ERG revised base case will apply the mid-point estimates.

DOC+ADT docetaxel treatment compliance estimate: mHSPC

As previously mentioned the docetaxel compliance estimates are only applied to the relatively minor £28 direct drug costs per docetaxel administration and not to the other cost elements such as chemotherapy administration costs. The mHSPC docetaxel compliance estimates are not applied in the same manner as the mHSPC abiraterone compliance estimates.

Compliance estimates: mCRPC

For mCRPC abiraterone and enzalutamide are assumed to be taken for 100% of the mCRPC discontinuation curve. These mCRPC treatments are mainly received in the ADT and DOC+ADT arms.

The other treatments have compliance percentages applied to them that do not take into account the effects of the MSM/TA387 model discontinuation curves; i.e., the compliance percentages of the trials will include some discontinuations. These mCRPC treatments are mainly received in the AAP+ADT arm.

Given the ERG's concerns around the mCRPC cost estimates the ERG have not attempted to address this issue.

Bone scans and CT scans

The company base case assumes that there will be no bone scans for AAP+ADT or for ADT, but that for DOC+ADT there will be and that the frequency of these will increase when patients have completed their course of docetaxel and are on ADT (post DOC+ADT)[q].

The ERG cannot find any reference to monitoring with bone scans or CT scans in either the docetaxel SmPC or the abiraterone SmPC, or any link from an increased risk of bone disease to this. Within the SmPCs it seems that LHRH agonists can reduce bone mineral density. The abiraterone SmPC states that "*Decreased bone density may occur in men with metastatic advanced prostate cancer. The use of ZYTIGA in combination with a glucocorticoid could increase this effect*". The ERG have not been able to find anything similar in the prostate cancer section of the docetaxel SmPC.

Given the above the ERG will equalise the number of bone scans in the DOC+ADT arm with those in the AAP+ADT arm. This has a reasonable impact upon the cost effectiveness of AAP+ADT compared to DOC+ADT, worsening it from £17,828 per QALY to £21,695 per QALY.

A scenario analysis will revert to the estimates of the company. Nevertheless, this does not particularly address:

- the frequency of bone scans for DOC+ADT for patients receiving docetaxel
- the frequency of bone scans for ADT (post DOC+ADT) patients in the longer term.

The second bullet is the more important.

[q] This differentiation of resource use between ADT (post DOC+ADT) and ADT also introduces some modelling complications if the costing of DOC+ADT takes into account compliance in the same manner as the costing of AAP+ADT.

Scenario analyses

The company do not provide any scenario analyses limiting the duration of treatment effect as outlined in the NICE methods guide section 5.1.16.

Calculation of mean time between rPFS and subsequent therapy: Minor issue

In response to the ERG clarification question B18 the company have confirmed that calculation of the mean time between rPFS and subsequent treatment is based upon all patients with rPFS data including those censored for time to subsequent therapy (TTST). At clarification the company has confirmed that there was a minor error in this calculation. Restricting the data to those with both an rPFS and a TTST event has a reasonable impact upon the estimates.

Table 62 Mean time between rPFS and subsequent therapy: months

	AAP+ADT	ADT	net
Patients with rPFS data			
Original submission	[redacted]	[redacted]	[redacted]
Correction at clarification	[redacted]	[redacted]	[redacted]
Patients with rPFS and TTST data	[redacted]	[redacted]	[redacted]

For the sake of argument suppose that all patients were recruited at the same time point with all patients in the ADT arm progressing at 6 and all patients in the AAP+ADT arm progressing at 7 months. Suppose further that the time between rPFS and subsequent therapy was 2 months in both arms and that IA1 corresponded to 8 months. The company method would estimate a mean time to treatment of 2 months in the ADT arm and 1 month in the AAP+ADT arm. While an extreme and unrealistic example, it does illustrate that for immature data the company method may be biased and underestimate the mean time from rPFS to subsequent treatment more for the arm that postpones rPFS for longer; i.e. in favour of AAP+ADT.

Similarly, given immature data, ignoring those censored for TTST may ignore those who never receive any subsequent treatment. This might bias the estimates in the opposite direction.

The estimates from the alternative method are quite different, though the net effect between the arms is less so. Applying them within the company model has little impact.

1st line mCRPC modelling during the 1st year of the MSM/TA387 model: Minor issue

The model needs to simulate newly incident mCRPC patients in every cycle of the model. This requires it to append the TA387 mCRPC discontinuation and OS curves to the newly incident mCRPC patients in each cycle. Error appears to have crept into the look-up of the cycle specific probabilities of discontinuation and death for mCRPC patients who are incident during the 1st year.

For instance, the week 1 incident mCRPC patients in the AAP+ADT arm have the correct weekly mortality probabilities applied up to week 52 of the model. At this point the model switches to a 4-weekly cycle and 4-weekly probabilities are applied. But rather than apply the 4-weekly probability from week 52 to week 56 of 4.6%, the model applies the 4-weekly probability from week 208 to week 212 of 7.2%. All the subsequent 4-weekly probabilities are similarly taken from 52*4=208 weeks too far down the relevant survival curves.

Application of full LATITUDE QoL regression for mCRPC: Minor Issue

If it is felt that the LATITUDE QoL regression should be applied in full when estimating the QoL values for the mHSPC health states, in the opinion of the ERG it should also be applied when deriving the QoL values for the mCRPC health states. The company scenario analysis around this only alters the QoL values for the mHSPC health states. However, applying parallel changes to the mCRPC health states has relatively little effect on the cost effectiveness estimates as it seems to affect all arms to much the same extent.

Abiraterone last administration cost: Minor issue

It is unclear why the company start dosing and costings not from baseline but from after 1 week. This may also be related to the fact that the company only apply 74% of the pack price of abiraterone for the last costed administration. This could in turn account for some of the differences in cost between the company model and the ERG rebuild.

5.4 Exploratory and sensitivity analyses undertaken by the ERG

A key question for the Committee for this appraisal is whether they prefer the MSM/TA387 model, the MSM model or is somewhere between the two. The ERG supply a full set of analyses for the MSM/TA387 model and a full set of analyses for the MSM model in what

follows. This mainly alters the balance between time spent on 1st line mCRPC treatment mCRPC and time spent on 3rd line mCRPC treatment.
The results of this section include the effects of the abiraterone commercial access arrangement but do not include the effects of the patient access schemes of enzalutamide, cabazitaxel or radium-223. These are supplied in the cPAS Appendix.

The ERG have revised the company model base case to:

- Apply the full set of LATITUDE quality of life regression coefficients, these also being rolled through to the quality of life values that are implied for mCRPC patients.
- Apply the LATITUDE quality of life regression that does not differentiate the SAE coefficient between the arms.
- Set the quality of life decrement for ADT (post DOC+ADT) relative to ADT to zero.
- Apply the compliance percentage for abiraterone derived from the CSR mid-point values.
- Apply compliance percentages in the DOC+ADT arm in the same manner as in the AAP+ADT arm.
- Apply the ERG preferred mCRPC treatment percentages that do not differentiate the proportions receiving BSC between the arms.
- Equalise the frequency of monitoring with bone and CT scans for those receiving a course of docetaxel in the DOC+ADT arm with those of the AAP+ADT arm.
- Equalise the frequency of monitoring with bone scans for those who have completed a course of docetaxel in the DOC+ADT arm, ADT (post DOC+ADT) patients, with those of the AAP+ADT arm.
- Apply corrections for minor issues.

Given the complexity of the company modelling the ERG provide a range of sensitivity and scenario analyses.

- SA01: Kaplan Meier to MSM TPM cut-offs of 4 months and of 7 months.
- SA02: Apply a common probability of PPS patients receiving 1st line mCRPC treatment for DOC+ADT and AAP+ADT, rather than conditioning the AAP+ADT probability of mCRPC treatment by the DOC+ADT hazard ratio for rPFS.
- SA03: Differentiate 1st line mCRPC treatment effectiveness in line with the central estimates of the company ITC.

- SA04: Differentiate 1st line mCRPC treatment effectiveness in line with the central estimates of the company ITC, also setting 1st line mCRPC abiraterone use to zero with these patients instead being treated with enzalutamide.
- SA05: Apply a quality of life increment for ADT (post DOC+ADT) compared to ADT of half that of AAP+ADT treatment effect of the LATITUDE quality of life regression.
- SA06: Apply a quality of life decrement for ADT (post DOC+ADT) compared to ADT of [redacted] as per the company base case.
- SA07: Drop the LATITUDE quality of life regression coefficients for SAEs and SREs and instead apply the smaller decrements that the company derives from the literature.
- SA08: Apply the LATITUDE quality of life regression that does differentiate the SAE coefficient between the arms.
- SA09: Apply the company base case mCRPC treatment percentages.
- SA10: Apply the compliance percentage for abiraterone derived by the company from the LATITUDE rPFS and TTD Kaplan Meier curves.
- SA11: Differentiate the frequency of monitoring with bone scans for those receiving a course of docetaxel and for those who have received a course of docetaxel in the past from that of the AAP+ADT arm as per the company base case.

The ERG revised base case which applies the MSM/TA387 model results in the following estimates.

Table 63 ERG revised base case: MSM/TA387 model: deterministic

	LYs	QALYs	Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	5.030	3.289	[redacted]				
ADT	3.505	2.213	[redacted]	1.525	1.076	£19,362	£17,992
DOCE	4.360	2.845	[redacted]	0.671	0.444	£13,965	£31,439

The MSM/TA387 model probabilistic estimates are aligned with the deterministic estimates. The CEACs are presented below.

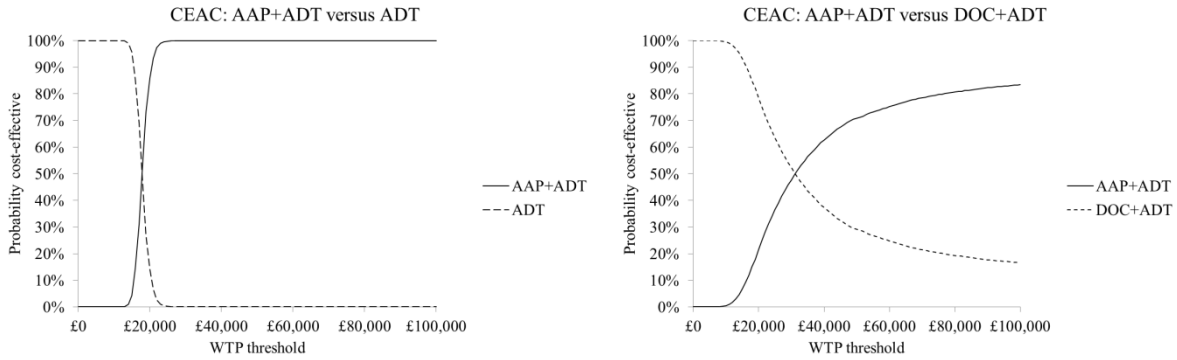


Figure 10 ERG revised base case: MSM/TA387 model: CEACs

The ERG revised base case which applies the MSM model results in the following estimates.

Table 64 ERG revised base case: MSM model: deterministic

	LYs	QALYs	Costs	Δ LYs	Δ QALYs	Δ Costs	ICER
AAP+ADT	5.129	3.249	[redacted]				
ADT	3.597	2.158	[redacted]	1.532	1.091	£22,751	£20,855
DOCE	4.365	2.761	[redacted]	0.764	0.488	£20,353	£41,697

The MSM model probabilistic estimates are aligned with the deterministic estimates. The CEACs are presented below.

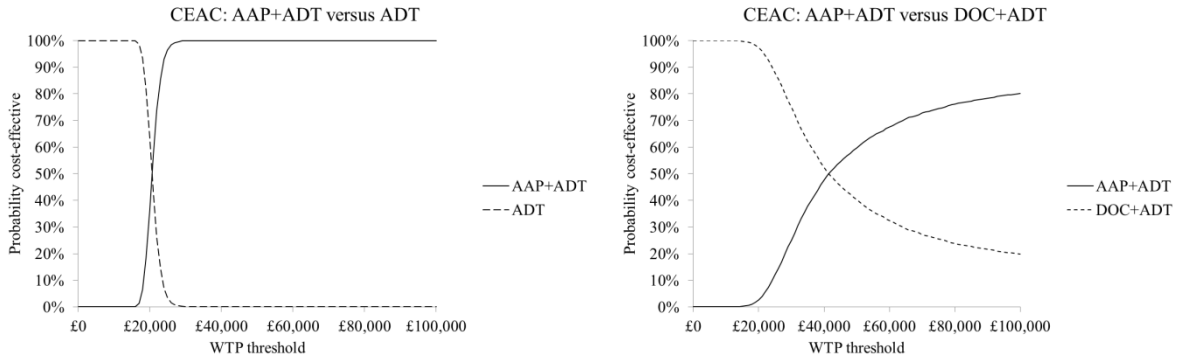


Figure 11 ERG revised base case: MSM model: CEACs

The ERG sensitivity analyses which apply the MSM/TA387 model result in the following estimates.

Table 65 ERG scenario analyses: MSM/TA387 model

	AAP+ADT vs ADT			AAP+ADT vs DOC+ADT		
	Δ QALYs	Δ Costs	ICER	Δ QALYs	Δ Costs	ICER
Base case	1.076	[redacted]	£17,992	0.444	[redacted]	£31,439
01a: KM 4mth	1.106	[redacted]	£17,479	0.460	[redacted]	£30,270
01b: KM 7mth	1.036	[redacted]	£18,453	0.419	[redacted]	£34,479
02: Same prob PPS Tx	..	[redacted]	..	0.441	[redacted]	£33,897
03: Diff effect mCRPC Tx	1.059	[redacted]	£17,687	0.425	[redacted]	£31,001
04: 03 + ENZA Tx prop.	1.049	[redacted]	£12,118	0.414	[redacted]	£16,714
05: DOC QoL increment	..	[redacted]	..	0.396	[redacted]	£35,255
06: DOC QoL decrement	..	[redacted]	..	0.516	[redacted]	£27,077
07: Company SAE/SRE QoL	1.112	[redacted]	£17,417	0.563	[redacted]	£24,805
08: Original LATITUDE QoL	1.086	[redacted]	£17,828	0.436	[redacted]	£32,046
09: Company mCRPC prop.	1.069	[redacted]	£18,336	0.437	[redacted]	£32,499
10: Company AAP % use	1.069	[redacted]	£16,837	0.437	[redacted]	£28,840
11: Company DOC scans	1.076	[redacted]	£18,181	0.444	[redacted]	£26,285

The ERG sensitivity analyses which apply the MSM model result in the following estimates.

Table 66 ERG scenario analyses: MSM model

	AAP+ADT vs ADT			AAP+ADT vs DOC+ADT		
	Δ QALYs	Δ Costs	ICER	Δ QALYs	Δ Costs	ICER
Base case	1.091	[redacted]	£20,855	0.488	[redacted]	£41,697
01a: KM 4mth	1.127	[redacted]	£20,295	0.503	[redacted]	£40,258
01b: KM 7mth	1.036	[redacted]	£21,407	0.462	[redacted]	£44,826
02: Same prob PPS Tx	..	[redacted]	..	0.483	[redacted]	£43,544
03: Diff effect mCRPC Tx	1.091	[redacted]	£20,858	0.488	[redacted]	£41,704
04: 03 + ENZA Tx prop.	1.093	[redacted]	£18,733	0.490	[redacted]	£37,562
05: DOC QoL increment	..	[redacted]	..	0.440	[redacted]	£46,253
06: DOC QoL decrement	..	[redacted]	..	0.560	[redacted]	£36,366
07: Company SAE/SRE QoL	1.127	[redacted]	£20,182	0.610	[redacted]	£33,386
08: Original LATITUDE QoL	1.101	[redacted]	£20,666	0.480	[redacted]	£42,425
09: Company mCRPC prop.	1.091	[redacted]	£21,690	0.488	[redacted]	£43,562
10: Company AAP % use	1.084	[redacted]	£19,735	0.481	[redacted]	£39,491
11: Company DOC scans	1.091	[redacted]	£20,903	0.488	[redacted]	£36,676

The two models are sensitive to the same elements:

- Applying the LATITUDE Kaplan Meier data for a longer period worsens the cost effectiveness estimates.
- Assuming that DOC+ADT patients who progress have the same probability of receiving treatment for mCRPC as those in the AAP+ADT arm worsens the cost effectiveness estimate.
- Differentiating 1st line mCRPC treatments' effectiveness has little effect. But assuming patients prefer enzalutamide rather than abiraterone for 1st line mCRPC treatment improves the cost effectiveness estimates. Both costs and QALYs are affected due to enzalutamide not being associated with a quality of life treatment effect increment compared to ADT, whereas abiraterone is.
- Quality of life increments and decrements for ADT (post DOC+ADT) have the predictable effects.
- Not applying the LATITUDE QoL regression in full but deriving SAE and SRE decrements from values in the literature improves the cost effectiveness estimates by quite a lot.
- Applying the company mHSPC abiraterone compliance percentage improves the cost effectiveness estimates.
- Applying the company bone scan frequencies for DOC+ADT improves the cost effectiveness estimates by quite a lot.

5.5 Conclusions of the cost effectiveness section

In the opinion of the ERG the 1st line mCRPC costs and benefits estimates of both the MSM/TA387 model and the MSM model are not reliable. This calls into question the reliability of the cost effectiveness estimates.

The company have a strong preference for the MSM/TA387 model over the MSM model. But given the ad hoc 2.62 OS hazard ratio, the implementation of the MSM/TA387 model is in large part an elaborate non-statistical method of fitting curves to the LATITUDE Kaplan Meier OS curves. This seems to negate the main company argument for developing the MSM/TA387 model: that neither the LATITUDE post rPFS survival data nor the LATITUDE Kaplan Meier OS curves are relevant to the UK. If curves are to be fitted to the LATITUDE Kaplan Meier OS curves it may be better to use the usual well-established

statistical methods, which would also allow time varying probabilities such as those of the Weibulls.

The company choose to apply the hazard ratio estimates of the mHSPC ITC to the AAP+ADT probabilities to estimate the DOC+ADT probabilities. It could equally well have chosen to apply them to the ADT probabilities to estimate the DOC+ADT probabilities. This worsens the cost effectiveness estimates. Both methods result in an anomalous estimate for DOC+ADT for the probability of mCRPC patients receiving 1st line mCRPC treatment.

The Committee may be more equipoise between the MSM model and the MSM/TA387 model than the company. The most important difference between them is the amount of time they model patients spending on 1st line, 2nd line and 3rd line mCRPC treatment. These durations are not affected by the ERG's concerns around the estimates of 1st line mCRPC costs and benefits. Alternatively, the Committee may prefer a partitioned survival analysis, or a presentation of fitted curves by way of model validation.

Given the company preferred modelling approach and the company ITC results [redacted]. This does not much affect results.

There is uncertainty about what 1st line mCRPC treatments proportions should be applied subsequent to AAP+ADT, ADT and DOC+ADT treatment for mHSPC. This is likely to become more important if the models' estimates of 1st line mCRPC treatments' costs and benefits are corrected.

The ERG view the company cost effectiveness estimates as perhaps biased in favour of AAP+ADT because:

- It is questionable whether there is a quality of life decrement for those who have completed a course of docetaxel compared to those who have only ever received ADT. There are reasons to suppose there may be an increment.
- If there is a quality of life decrement for those who have completed a course of docetaxel the company commissioned TTO study that estimates this may be biased.

- The company estimates of the quality of life decrements for docetaxel are only applied in the DOC+ADT arm, and not to docetaxel treatment for mCRPC in the AAP+ADT arm or the ADT arm.
- The company only partially apply the results of the LATITUDE QoL regression, which pushes up quality of life values to above those observed during the LATITUDE trial.
- The treatment compliance estimate for abiraterone for mHSPC seems low compared to CSR data on compliance.
- The treatment compliance estimate for docetaxel for mHSPC is not applied to the same range of costs as the compliance estimate for abiraterone.
- The treatment compliance estimates for mCRPC do not take into account that they reflect discontinuations during the relevant trials. Unadjusted compliance rates are applied to the MSM/TA387 model mCRPC treatment discontinuation curves. This mainly affects mCRPC treatments in the AAP+ADT arm.
- The ERG cannot find evidence that DOC+ADT is associated with more bone scans than both AAP+ADT and ADT, or that ADT (post DOC+ADT) is associated with more bone and CT scans than both AAP+ADT and ADT. It is mainly the latter that affects results.

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The ERG have made a number of revisions and corrections to the MSM/TA387 model and the MSM model. Most notably:

- Applying the full LATITUDE quality of life regression so that the quality of life values reflect those observed during the trial.
- Not applying the company quality of life decrement for those who have completed a course of docetaxel for mHSPC. The ERG consider the evidence presented by the company for this as rather thin. There is trial data which suggest there may actually be an increment.
- Applying an estimate of the proportion of abiraterone mHSPC patients that incurs abiraterone treatment costs based upon compliance data in the clinical study report. The company estimate derived from the LATITUDE rPFS and TTD curves seems too low, particularly towards the end of these curves.
- Equalising the frequency of bone scans for those who have completed a course of docetaxel for mHSPC with those receiving abiraterone for mHSPC in the AAP+ADT arm.

Each of these changes has a reasonable impact upon the cost effectiveness estimates. The full details of this and other sensitivity analyses are presented in section 5.4 above.

When using the MSM/TA387 model these changes taken together worsen the cost effectiveness estimates from £17,418 per QALY to £17,992 per QALY for the comparison of AAP+ADT with ADT and from £17,828 per QALY to £31,439 per QALY for the comparison of AAP+ADT with DOC+ADT.

When using the MSM model these changes taken together worsen the cost effectiveness estimates from £20,438 per QALY to £20,855 per QALY for the comparison of AAP+ADT with ADT and from £26,909 per QALY to £41,697 per QALY for the comparison of AAP+ADT with DOC+ADT.

The probabilistic estimates are aligned with these deterministic estimates.

7 Overall conclusions

The company's submission considered abiraterone acetate (Zytiga, Janssen-Cilag Ltd.) with prednisone/prednisolone (AAP) plus androgen deprivation therapy (ADT) for the treatment of adults with newly diagnosed, high risk mHSPC.

7.1 Clinical effectiveness evidence

The NICE final scope specified AAP+ADT compared with ADT alone or docetaxel+ADT in adults with newly diagnosed high risk metastatic hormone-naïve prostate cancer (mHNPC). The population addressed in the company submission is adults with newly diagnosed, high risk mHSPC. The company state that the terms mHNPC and newly diagnosed mHSPC are effectively the same because newly diagnosed patients are, by default, hormone naïve. The company submission also did not consider orchidectomy and bicalutamide monotherapy, as part of ADT alone treatment, as their clinical experts advised that these are seldom used in the UK.

The submission focuses on the results of the LATITUDE trial, which provide evidence of the benefits of AAP over ADT for the treatment of men with mHSPC. The benefit found in LATITUDE is evident for the primary outcomes of overall survival and progression measured by rPFS and extends to the secondary outcomes for safety and quality of life. The results of LATITUDE are similar to those from the STAMPEDE study. However, the STAMPEDE patient group was broader and while the company have conducted similar analyses on a post hoc subgroup meant to be similar to the LATITUDE population, they rightly have not combined them in any further analyses.

Less reliable are the company results of AAP compared with other treatments, predominately docetaxel. With no head-to-head studies available, these were compared using indirect methods. The company chose NMA at this stage, which the ERG agree, was sensible. When conducting the NMA the company used the recommended WinBUGS program from the NICE DSU TSD 2.[49] They were restricted to only fixed effects models because of the lack of studies and links between treatment groups. Further concerns are the many aspects of heterogeneity between the

studies, all recognised by the company. So while the ERG confirm the results provided showing AAP to be at least equivalent to docetaxel, there is a concern that estimates from these results will not be robust. There were no checks of statistical heterogeneity or consistency commented on. As such any economic modelling on these estimates will require caution and various scenarios to reflect these concerns.

The company also attempted to assess the efficacy of AAP+ADT for patient with disease progression (mCRPC) compared with other treatments. They focus on docetaxel, radium-223 and enzalutamide. The more robust method of NMA was not chosen and instead the company used Bucher pairwise comparisons. While NMA are more useful when making choices between multiple alternatives, the ERG confirm that NMA models did not converge probably due to the limited number of studies and data so that Bucher pairwise estimates were a reasonable alternative. For this patient group, the estimates show AAP to be comparable with other treatments. However, since checks of statistical heterogeneity or fit were not provided and as before the conceptual heterogeneity (e.g., differences in study populations, study setting, follow-up procedures, outcome measures) was extensive, caution for further economic modelling is warranted.

7.2 Cost-effectiveness evidence

It appears that the 1st line mCRPC costs and benefits estimates of both the MSM/TA387 model and the MSM model are not reliable. All cost effectiveness estimates may consequently not be reliable.

The company have a strong preference for the MSM/TA387 model over the MSM model. Due to the ad hoc 2.62 hazard ratio this is in large part an elaborate non-statistical method of fitting curves to the LATITUDE Kaplan Meier OS curves. If curves are to be fitted to the LATITUDE Kaplan Meier OS curves it may be better to use the usual well-established statistical methods, which would also allow time varying probabilities to be explored.

The Committee for this appraisal may be more equipoise between the MSM model and the MSM/TA387 model than the company. The most important difference between them is the amount of time they model patients spending on 1st line, 2nd line

and 3rd line mCRPC treatment. Alternatively, the Committee may prefer a partitioned survival analysis, or a presentation of statistically fitted curves by way of model validation.

There is uncertainty about what 1st line mCRPC treatments proportions should be applied subsequent to AAP+ADT, ADT and DOC+ADT treatment for mHSPC. This is likely to become more important if the models' estimates of 1st line mCRPC treatments' costs and benefits are corrected.

The company cost effectiveness estimates may be biased in favour of AAP+ADT because:

- It is questionable whether there is a quality of life decrement for those who have completed a course of docetaxel compared to those who have only ever received ADT. There are reasons and trial data to suppose there may be an increment.
- If there is a quality of life decrement for those who have completed a course of docetaxel the company commissioned TTO study that estimates this may be biased.
- The company estimates of the quality of life decrements for docetaxel are only applied in the DOC+ADT arm, and not to docetaxel treatment for mCRPC in the AAP+ADT arm or the ADT arm.
- The company only partially apply the results of the LATITUDE QoL regression, which pushes up quality of life values to above those observed during the LATITUDE trial.
- The treatment compliance estimate for abiraterone for mHSPC seems low compared to the CSR data on compliance.
- The treatment compliance estimate for docetaxel for mHSPC is not applied to the same range of costs as the compliance estimate for abiraterone.
- The treatment compliance estimates for mCRPC do not take into account that they reflect discontinuations during the relevant trials. This mainly affects mCRPC treatments in the AAP+ADT arm.
- The ERG cannot find evidence that mHSPC patients who have completed their course of docetaxel and are only receiving ADT in the DOC+ADT arm have more routine bone scans than mHSPC patients in the AAP+ADT arm.

8 References

1. *Prostate cancer incidence statistics*: Cancer Research UK; 2017. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence [Accessed April 2018]

2. Frees S, Akamatsu S, Lynch K, Chavez-Munoz C, Black P, Gleave M, et al. Timing of the development of metastasis and initiation of treatment are important prognostic factors in prostate cancer. *Eur Urol Suppl* 2016;**15 (3):**e845.

3. Sweeney C, Chen YH, Liu G, Carducci M, Jarrard D, Eisenberger M, et al. Long term efficacy and QOL data of chemohormonal therapy (C-HT) in low and high volume hormone naive metastatic prostate cancer (PrCa): E3805 CHAARTED trial. *Ann Oncol* 2016;**27**.

4. *Prostate cancer survival statistics*: Cancer Research UK; 2017. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/survival [Accessed April 2018]

5. *Prostate cancer: stages and grades*: Cancer.net; 2017. https://www.cancer.net/cancer-types/prostate-cancer/stages-and-grades [Accessed April 2018]

6. Koo KC, Park SU, Kim KH, Rha KH, Hong SJ, Yang SC, et al. Prognostic impacts of metastatic site and pain on progression to castrate resistance and mortality in patients with metastatic prostate cancer. *Yonsei Med J* 2015;**56:**1206-12.

7. Mottet N, De Santis M, Briers E, Bourke L, Gillessen S, Grummet JP, et al. Updated Guidelines for Metastatic Hormone-sensitive Prostate Cancer: Abiraterone Acetate Combined with Castration Is Another Standard *Eur Urol* 2018;**73:**316-21.

8. Buelens S, Poelaert F, Dhondt B, Fonteyne V, De Visschere P, Ost P, et al. Metastatic burden in newly diagnosed hormone-naive metastatic prostate cancer: Comparing definitions of CHAARTED and LATITUDE trial. *Urol Oncol* 2018.

9. Albertsen PC, Aaronson NK, Muller MJ, Keller SD, Ware Jr JE. Health-related quality of life among patients with metastatic prostate cancer. *Urology* 1997;**49:**207-17.

10. Krahn MD, Bremner KE, Zagorski B, Alibhai SM, Chen W, Tomlinson G, et al. Health care costs for state transition models in prostate cancer. *Med Decis Making* 2014;**34:**366-78.

11. Oster G, Lamerato L, Glass AG, Richert-Boe KE, Lopez A, Chung K, et al. Natural history of skeletal-related events in patients with breast, lung, or prostate cancer and metastases to bone: A 15-year study in two large US health systems. *Support Care Cancer* 2013;**21:**3279-86.

12. James N, Pirrie S, Pope A, Barton D, Andronis L, Goranitis I, et al. TRAPEZE: A randomised controlled trial of the clinical effectiveness and cost-effectiveness of chemotherapy with zoledronic acid, strontium-89, or both, in men with bony metastatic castration-refractory prostate cancer. *Health Technol Assess* 2016;**20:**1-127.

13. Eton DT, Lepore SJ. Prostate cancer and health-related quality of life: A review of the literature. *Psychooncology* 2002;**11:**307-26.

14. Bernard B, Sweeney CJ. Management of Metastatic Hormone-Sensitive Prostate Cancer. *Curr Urol Rep* 2015;**16**.

15. *Prostate cancer. NCCN Clinical Practice Guidelines on Oncology*: National Comprehensive Cancer Network; 2017. https://www.nccn.org/professionals/physician_gls/pdf/prostate_blocks.pdf [Accessed April 2018]

16. *Prostate cancer guidelines*: European Association of Urology; 2018. http://uroweb.org/guideline/prostate-cancer/ [Accessed April 2018]

17. Rulach RJ, McKay S, Neilson S, White L, Wallace J, Carruthers R, et al. Real-world uptake, safety profile and outcomes of docetaxel in newly diagnosed metastatic prostate cancer. *BJU Int* 2018;**121:**268-74.

18. Parker C, Gillessen S, Heidenreich A, Horwich A. Cancer of the prostate: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015;**26:**v69-v77.

19. Aly A, Mullins CD, Hussain A. Understanding heterogeneity of treatment effect in prostate cancer. *Curr Opin Oncol* 2015;**27:**209-16.

20. James ND, Spears MR, Clarke NW, Dearnaley DP, De Bono JS, Gale J, et al. Survival with newly diagnosed metastatic prostate cancer in the docetaxel era: Data from 917 patients in the control arm of the STAMPEDE Trial (MRC PR08, CRUK/06/019). *Eur Urol* 2015;**67:**1028-38.

21. Sweeney CJ, Chen YH, Carducci M, Liu G, Jarrard DF, Eisenberger M, et al. Chemohormonal therapy in metastatic hormone-sensitive prostate cancer. *N Engl J Med* 2015;**373:**737-46.

22. Pharma A. *Taxotere: Summary of Product Characteristics*: European Medicine Agency. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000073/WC500035264.pdf [Accessed April 2018]

23. Al-Batran SE, Hozaeel W, Tauchert FK, Hofheinz RD, Hinke A, Windemuth-Kieselbach C, et al. The impact of docetaxel-related toxicities on health-related quality of life in patients with metastatic cancer (QoliTax). *Ann Oncol* 2015;**26:**1244-8.

24. Ito T, Grant L, Mills A, Heselwood A, Gater A. Qualitative exploration of the real-world experiences of men receiving or refusing chemotherapy (docetaxel) for the treatment of metastatic hormone-sensitive prostate cancer (MHSPC). *Value Health* 2017;**20 (9):**A450.

25. Patrikidou A, Uccello M, Tree A, Parker C, Attard G, Eeles R, et al. Upfront Docetaxel in the Post-STAMPEDE World: Lessons from an Early Evaluation of Non-trial Usage in Hormone-Sensitive Prostate Cancer. *Clin Oncol* 2017;**29:**e174-e5.

26. Cornford P, Bellmunt J, Bolla M, Briers E, De Santis M, Gross T, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part II: Treatment of Relapsing, Metastatic, and Castration-Resistant Prostate Cancer. *Eur Urol* 2017;**71:**630-42.

27. Fizazi K, Jenkins C, Tannock IF. Should docetaxel be standard of care for patients with metastatic hormone-sensitive prostate cancer? Pro and contra. *Ann Oncol* 2015;**26:**1660-7.

28. Gravis G, Boher JM, Joly F, Soulie M, Albiges L, Priou F, et al. Androgen Deprivation Therapy (ADT) Plus Docetaxel Versus ADT Alone in Metastatic Non castrate Prostate Cancer: Impact of Metastatic Burden and Long-term Survival Analysis of the Randomized Phase 3 GETUG-AFU15 Trial. *Eur Urol* 2016;**70:**256-62.

29. *Prostate cancer: diagnosis and management. Clinical guideline CG175*. London: National Institute for Health and Care Excellence; 2014. https://www.nice.org.uk/guidance/cg175 [Accessed April 2018]

30. *Abiraterone for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen. Technology appraisal guidance TA259*. London: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ta259 [Accessed April 2018]

31. *Abiraterone for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated. Technology appraisal guidance TA387*. London: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ta387 [Accessed April 2018]

32. *Docetaxel for the treatment of hormone-refractory metastatic prostate cancer. Technology appraisal guidance TA101*. London: National Institute for Health and Care Excellence; 2006. https://www.nice.org.uk/guidance/ta101 [Accessed April 2018]

33. *Enzalutamide for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated. Technology appraisal guidance TA377*. London: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ta377 [Accessed April 2018]

34. *Enzalutamide for metastatic hormone‐relapsed prostate cancer previously treated with a docetaxel‐containing regimen. Technology appraisal guidance TA316*. London: National Institute for Health and Care Excellence; 2014. https://www.nice.org.uk/guidance/ta316 [Accessed April 2018]

35. *Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel. Technology appraisal guidance TA391*. London: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ta391 [Accessed April 2018]

36. *Radium-223 dichloride for treating hormone-relapsed prostate cancer with bone metastases. technology appraisal guidance TA412*. London: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ta412 [Accessed April 2018]

37. *Zytiga (abiraterone acetate). European public assessment report*. London: European Medicines Agency; 2017. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/002321/WC500112861.pdf [Accessed April 2018]

38. NV J-CI. *Zytiga Summary of product characteristics*. London: European Medicines Agency; 2017. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002321/WC500112858.pdf [Accessed April 2018]

39. Chen Y, Clegg NJ, Scher HI. Anti-androgens and androgen-depleting therapies in prostate cancer: new agents for an established target. *Lancet Oncol* 2009;**10:**981-91.

40. De Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med* 2011;**364:**1995-2005.

41. Fizazi K, Tran N, Fein L, Matsubara N, Rodriguez-Antolin A, Alekseev BY, et al. Abiraterone plus prednisone in metastatic, castration-sensitive prostate cancer. *N Engl J Med* 2017;**377:**352-60.

42. Chi KN, Protheroe A, Rodriguez-Antolin A, Facchini G, Suttman H, Matsubara N, et al. Patient-reported outcomes following abiraterone acetate plus prednisone added to androgen deprivation therapy in patients with newly diagnosed metastatic castration-naive prostate cancer (LATITUDE): An international, randomised phase 3 trial. *Lancet Oncology* 2018.

43. Feyerabend S, Saad F, Li T, Ito T, Diels J, Van Sanden S. Indirect treatment comparison (ITC) of abiraterone acetate (AA) plus prednisone (P) and docetaxel (DOC) on patient-reported outcomes (PROs) in metastatic castration-naïve prostate cancer (mCNPC). *J Clin Oncol* 2018;**36:**2000.

44. Fizazi K, Tran N, Fein L, Matsubara N, Bjartell A. Abiraterone acetate (AA) plus prednisone (P) 5 mg QD in metastatic castration-naïve prostate cancer (mCNPC): Detailed safety analyses from the LATITUDE phase 3 trial. . *J Clin Oncol* 2018;**36:**182.

45. James ND, De Bono JS, Spears MR, Clarke NW, Mason MD, Dearnaley DP, et al. Abiraterone for prostate cancer not previously treated with hormone therapy. *N Engl J Med* 2017;**377:**338-51.

46. *Systematic reviews: CRD's guidance for undertaking systematic reviews in health care [document on the Internet]*. University of York 2009. URL: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm [Accessed April 2018]

47. Sydes MR, Mason MD, Spears MR, Clarke NW, Dearnaley D, Ritchie AWS, et al. PR Adding abiraterone acetate plus prednisolone (AAP) or docetaxel for patients (pts) with high-risk prostate cancer (PCa) starting long-term androgen deprivation therapy (ADT): Directly randomised data from STAMPEDE (NCT00268476). *Ann Oncol* 2017;**28 (Supplement 5):**v619.

48. Rydzewska LHM, Burdett S, Vale CL, Clarke NW, Fizazi K, Kheoh T, et al. Adding abiraterone to androgen deprivation therapy in men with metastatic hormone-sensitive prostate cancer: A systematic review and meta-analysis. *Eur J Cancer* 2017;**84:**88-101.

49. Dias S, Welton N, Sutton A, Ades A. *A generalised linear modelling framework for pairwise and netwrok meta-analysis of randomised controlled trials. NUCE DSU tecnical support document 2*. Sheffield: NICE Decision Support Unit; 2016. http://nicedsu.org.uk/wp-content/uploads/2017/05/TSD2-General-meta-analysis-corrected-2Sep2016v2.pdf [Accessed April 2018]

50. Vale CL, Fisher DJ, Carpenter J, White IR, Burdett S, Clarke NW, et al. What are the optimal systemic treatments for men with metastatic, hormone-sensitive prostate cancer? A STOPCaP systematic review and network meta-analysis. *Ann Oncol* 2017;**28 (Supplement 5):**v620.

51. Wallis CJD, Klaassen Z, Bhindi B, Goldberg H, Chandrasekar T, Farrell AM, et al. Comparison of Abiraterone Acetate and Docetaxel with Androgen Deprivation Therapy in High-risk and Metastatic Hormone-naive Prostate Cancer: A Systematic Review and Network Meta-analysis. *Eur Urol* 2017;**13:**https://doi.org/10.1016/j.eururo.2017.10.002.

52. Joly F, Gravis G, Oudard S, Soulie M, Esterni B, Protiere C, et al. Patients' self-assessment of tolerance and quality of life during docetaxel-based chemotherapy: Results from a phase III trial in hormone-naive metastatic prostate cancer patients (GETUG-AFU 15/0403). *J Clin Oncol* 2010;**28:**28.

53. Patrick-Miller LJ, Chen YH, Carducci MA, Cella D, DiPaola RS, Gartrell BA, et al. Quality of life (QOL) analysis from CHAARTED: Chemohormonal androgen ablation randomized trial in prostate cancer (E3805). *J Clin Oncol* 2016;**34:**26.

54. Morgans AK, Chen YH, Sweeney CJ, Jarrard DF, Plimack ER, Gartrell BA, et al. Quality of Life During Treatment With Chemohormonal Therapy: Analysis of E3805 Chemohormonal Androgen Ablation Randomized Trial in Prostate Cancer. *J Clin Oncol* 2018JCO2017753335.

55. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: An online database and application of the MAPS statement. *Health Qual Life Outcome* 2018;**16** 31.

56. Van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15:**708-15.

57. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): Survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet* 2016;**387:**1163-77.

58. Gravis G, Fizazi K, Joly F, Oudard S, Priou F, Esterni B, et al. Androgen-deprivation therapy alone or with docetaxel in non-castrate metastatic prostate cancer (GETUG-AFU 15): A randomised, Open-label, Phase 3 trial. *Lancet Oncol* 2013;**14:**149-58.

59. Hernandez Alava M, Wailoo A, Pudney S. *Methods for mapping between the EQ-5D-5L and the 3L for technology appraisal*. Sheffield: NICE Decision Support

Unit; 2017. http://nicedsu.org.uk/wp-content/uploads/2017/05/Mapping-5L-to-3L-DSU-report.pdf [Accessed April 2018]

60. *NICE impact; Cancer*: National Institute for Health and Care Excellence; 2018. https://www.nice.org.uk/media/default/about/what-we-do/into-practice/measuring-uptake/nice-impact-cancer.pdf [Accessed April 2018]

9 Appendices

Appendix 1 TTO study three main health states

ADT	DOC+ADT	ADT (post DOC+ADT)
[redacted]		