# University of Groningen

## Abiraterone Acetate for the Treatment of Chemotherapy-Naive Metastatic Castration-Resistant Prostate Cancer

Ramaekers, Bram L. T.; Riemsma, Rob; Tomini, Florian; van Asselt, Thea; Deshpande, Sohan; Duffy, Steven; Armstrong, Nigel; Severens, Johan L.; Kleijnen, Jos; Joore, Manuela A.

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

CrossMark

REVIEW ARTICLE

# Abiraterone Acetate for the Treatment of Chemotherapy-Naïve Metastatic Castration-Resistant Prostate Cancer: An Evidence Review Group Perspective of an NICE Single Technology Appraisal

Bram L. T. Ramaekers[1] · Rob Riemsma[2] · Florian Tomini[1] · Thea van Asselt[1,3,4] ·
Sohan Deshpande[2] · Steven Duffy[2] · Nigel Armstrong[2] · Johan L. Severens[5] ·
Jos Kleijnen[2,6] · Manuela A. Joore[1]

**Abstract** The National Institute for Health and Care Excellence (NICE) invited Janssen, the company manufacturing abiraterone acetate (AA; tradename Zytiga®), to submit evidence for the clinical and cost effectiveness of AA in combination with prednisone/prednisolone (AAP) compared with watchful waiting (i.e. best supportive care [BSC]) for chemotherapy-naïve patients with metastatic castration-resistant prostate cancer (mCRPC). Kleijnen Systematic Reviews Ltd (KSR), in collaboration with Maastricht University Medical Center, was commissioned as the Evidence Review Group (ERG). This paper presents a summary of the company submission (CS), the ERG report, subsequent addenda, and the development of the NICE guidance for the use of this drug in England and Wales by the Appraisal Committee (AC). The ERG produced a critical review of the clinical and cost effectiveness of AAP based on the CS. An important question in this appraisal was, according to the ERG, whether AAP followed by docetaxel is more effective than BSC followed by docetaxel. In the COU-AA-302 trial, 239 of 546 (43.8 %) AAP patients and 304 of 542 (56.1 %) BSC patients received docetaxel as subsequent therapy, following AA or placebo. The results for this specific group of patients were not presented in the CS; therefore, the ERG asked the company to provide these data in the clarification letter; however, these data were presented as commercial-in-confidence and cannot therefore be reported here. The ERG's critical assessment of the company's economic evaluation highlighted a number of concerns, including (a) not using the intention-to-treat (ITT) population; (b) inconsistencies in estimating prediction equations; (c) not fully incorporating the impact of adverse events; (d) incorrectly incorporating the new patient access scheme (PAS); and (e) the assumption that AA non-compliance leads to recoverable drug costs. Although some of these issues were adjusted in the ERG base case, the ERG could not estimate the impact of all of these issues, and thus acknowledges that there are still uncertainties concerning the cost-effectiveness evidence. With the exception of the ERG's preference for using the ITT population, the AC agreed with the approach taken in the ERG base case. The original company and ERG base-case incremental cost-effectiveness ratios (ICERs) were £46,722 and £57,688 per QALY gained, respectively; these changed to £28,563 and £38,061 per QALY gained, respectively, in the revised base cases applying a new PAS. Regarding the end-of-life criteria, after 24 months approximately 63 % of patients in the control group of the COU-AA-302 trial were still alive, and the median survival was 30.1 months (95 % CI 27.3–34.1). Therefore, it is unlikely that life expectancy would be less than 24 months. The AC stated that the most plausible ICER is likely between £28,600 and £32,800 per

✉ Bram L. T. Ramaekers
  bram.ramaekers@mumc.nl

[1] Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Center, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands

[2] Kleijnen Systematic Reviews Ltd, York, UK

[3] Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[4] Department of Pharmacy, University of Groningen, Groningen, The Netherlands

[5] Institute of Health Policy and Management, iMTA-Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands

[6] School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands

QALY gained, and concluded that AAP at this stage in the treatment pathway did not meet the end-of-life criterion for short life expectancy. Moreover, in March 2016, the AC produced the final guidance, stating that AAP is recommended, within its marketing authorisation, as an option for treating mCRPC.

---

**Key Points for Decision Makers**

Abiraterone acetate (tradename Zytiga®) in combination with prednisone/prednisolone (AAP) delays clinical disease progression and initiation of chemotherapy compared with best supportive care in chemotherapy-naïve patients with metastatic castration-resistant prostate cancer (mCRPC).

Typically, the ITT population is preferred to populate the economic model; however, in this specific case, the National Institute for Health and Care Excellence (NICE) Appraisal Committee preferred a selected subpopulation (complete cases).

Multiple patient access schemes (PASs) for the same drug might be used in the economic model if the drug is used in different disease stages.

Potential administration costs of complex PASs should be incorporated in the cost-effectiveness estimates.

The NICE Appraisal Committee has recommended AAP within its licenced indication as an option for treating mCRPC in people who have no or mild symptoms after androgen deprivation therapy has failed, and before chemotherapy is indicated.

---

## 1 Introduction

Health technologies must be shown to be clinically effective and to represent a cost effective use of National Health Service (NHS) resources in order to be recommended by the National Institute for Health and Care Excellence (NICE) for use within the NHS in England and Wales. NICE is an independent organisation responsible for providing national guidance on promoting good health and preventing and treating ill health in priority areas with a significant impact. The NICE single technology appraisal (STA) process typically considers new technologies within a single indication [1]. Within the STA process, the company provides NICE with a written submission, including

an executable health economic model, considering the company's estimates of the clinical effectiveness and cost effectiveness of the technology. This company submission (CS) is critically reviewed by the Evidence Review Group (ERG), an external organisation independent of NICE, which produces an ERG report. After consideration of the CS, the ERG report, and testimony from experts and other stakeholders, the NICE Appraisal Committee (AC) formulates the Appraisal Consultation Document (ACD), which contains preliminary guidance regarding the initial decision on whether or not to recommend the technology. Subsequently, stakeholders are invited to comment on the submitted evidence and the ACD, after which a subsequent ACD may be produced or a Final Appraisal Determination (FAD) is issued, which is open to appeal.

This paper presents a summary of the CS [2], the ERG report [3], subsequent addenda [4] for the STA on abiraterone acetate (AA; tradename: Zytiga®) in combination with prednisone/prednisolone (AAP) for the treatment of chemotherapy-naïve metastatic castration-resistant prostate cancer (mCRPC), and the subsequent development of the NICE guidance. All relevant documents are publically available online [4]. AAP has previously been recommended by NICE for the treatment of mCRPC previously treated with docetaxel-containing chemotherapy (STA259) [5].

## 2 The Decision Problem

The patient population described in the final scope, specified by NICE [6], is "Adults with mCRPC who are asymptomatic or mildly symptomatic after failure of androgen deprivation therapy (ADT) in whom chemotherapy is not yet clinically indicated".

In the early stages, prostate cancer is localised to the prostate gland and can be managed with active surveillance, surgical removal of the prostate (i.e. prostatectomy) or radical radiotherapy with or without ADT [7]. However, it may slowly progress to a chronic stage and over a period of time can rapidly progress to a more advanced/metastatic stage [2]. It is estimated that 55–65 % of prostate cancer patients will develop metastatic disease [6]. Available treatments for metastatic prostate cancer include surgical castration or ADT to reduce the testosterone levels, which helps in slowing down the tumour growth and delays progression. Nevertheless, after 1–2 years the tumour typically stops responding to the castration therapy and resumes growth [2]; this is termed 'castration-resistant' prostate cancer. The patients diagnosed with 'castration-resistant' prostate cancer are likely to be metastatic (i.e. mCRPC), meaning the tumour has spread outside the prostate. According to the CS [2], it

was traditionally thought that tumours grow during ADT because they became 'hormone-refractory' or 'androgen-independent'. However, current knowledge indicates that these tumours still rely on hormones such as testosterone for their growth, but are dependent on alternative sources (e.g. adrenal cortex and synthesis within the tumour itself) [2, 8]. For mCRPC, docetaxel is recommended as a treatment option for hormone-refractory prostate cancer associated with a Karnofsky performance status score of 60 % or more [7].

The company stated that the most common complaints reported by symptomatic mCRPC patients included lower extremity pain, loss of appetite and weight loss, skeletal-related events (SREs), renal failure due to obstruction of the urethra, and oedema due to obstruction of venous and lymphatic tributaries by nodal metastases [2, 9, 10].

When converted in vivo to abiraterone, AA is a selective androgen biosynthesis that blocks cytochrome P17 (17α-hydroxylase; an enzyme thought to play a role in the production of testosterone), thereby stopping the testes and other tissues in the body from making testosterone. Treatment with AA therefore decreases serum testosterone to undetectable levels, while ADT, such as luteinising hormone-releasing hormone analogues, decrease androgen production in the testes but do not affect androgen production by the adrenals or in the tumour [2].

In December 2012, AA received a marketing authorisation from the European Medicines Agency (EMA) for the treatment of mCRPC in adult men who are asymptomatic or mildly symptomatic after failure of ADT in whom chemotherapy is not yet clinically indicated. The recommended dose is 1000 mg (single daily dose administered orally) in combination with low-dose prednisone/prednisolone (recommended dose 10 mg daily). The most common adverse reactions seen are peripheral oedema, hypokalaemia, hypertension and urinary tract infection. Other important adverse reactions include cardiac disorders, hepatotoxicity, fractures, and allergic alveolitis [11].

NICE issued a final scope [6] in January 2013 to appraise the clinical and cost effectiveness of AAP within its licensed indication for the treatment of chemotherapy-naïve mCRPC. At the time of submission, the scope stated that the current relevant treatment options within the NHS include docetaxel and best supportive care (BSC; may include radiotherapy, radiopharmaceuticals, analgesics, bisphosphonates, further hormonal therapies, and corticosteroids). Other subsequently licensed treatment options were not considered relevant treatment options in this STA (e.g. enzalutamide [tradename Xtandi®], which is now recommended [12]).

# 3 Independent Evidence Review Group Review (ERG)

In February 2014, the company (Janssen) provided a submission to NICE on the clinical and cost effectiveness of AAP within its licensed indication. In conformity with the process for STAs, the company provided additional information in response to clarification questions raised by the ERG and NICE. Additionally, the ERG adjusted the decision analytic model received from the company to assess the impact of alternative parameter values and assumptions on the model results and to produce an ERG base case. Sections 3.1–3.4 below summarises the evidence presented in the CS, as well as the ERG's review of that evidence. Moreover, four addenda [4] submitted by the ERG (upon request) in response to questions raised by the AC, to additional data provided by the company and to a new patient access scheme (PAS) submitted by the company, are discussed. PASs typically reflect a discount on the list price of a drug and are designed to ensure patients can gain access to high-cost drugs.

## 3.1 Clinical-Effectiveness Evidence Submitted by the Company

One randomised controlled trial (the COU-AA-302 trial [13–15]) was included for the comparison of AAP versus BSC. In the COU-AA-302 trial, a total of 1088 patients were recruited and randomised to AAP ($n = 546$) or BSC (i.e. placebo plus prednisone/prednisolone [PP]; $n = 542$). Overall, 1082 patients received at least one dose of the allocated intervention (safety population). Patients continued treatment with AAP or BSC until disease progression (determined according to radiographic and clinical measures). The median treatment duration was 13.8 months (15 cycles initiated) in the AAP arm, and 8.3 months (nine cycles initiated) in the BSC arm.

Results presented in the CS [2] were based on the results from the second (data cut-off 20 December 2011) and third (data cut-off 22 May 2012) interim analyses of the COU-AA-302 study [13–15], which were conducted after approximately 40 and 55 % of the total overall survival (OS) events had occurred. Neither the second nor third interim analysis OS results met the prespecified statistical significance levels (hazard ratio [HR] at third interim analysis 0.79; 95 % CI 0.66–0.96). Median OS was 35.3 months (95 % CI 31.2–35.3) in the AAP group and 30.1 months (95 % CI 27.3–34.1) in the PP group. The company did not provide mean survival for both groups or mean survival gain, despite explicit questions in the clarification letter by the ERG [16].

Treatment with AAP resulted in a 48 % relative reduction in the risk of radiographic progression compared with PP (absolute risk reduction 11.5 %), and increased progression-free survival by 8.2 months. Significant differences in favour of the AAP group were observed for objective response rate (complete or partial response according to modified Response Evaluation Criterita in Solid Tumors [RECIST] criteria), prostate-specific antigen response, and duration of response. Health-related quality of life (HRQoL) was assessed in the COU-AA-302 study via the Functional Assessment of Cancer Therapy–Prostate (FACT–P) instrument; however, no results were reported by treatment arm for baseline, follow-up, or change scores. Time to progression in average pain intensity and worst pain intensity showed no significant differences between treatment arms. All other pain-related outcomes favoured AAP over BSC.

Adverse events (AEs) were significantly more often reported in the AAP arm when compared with the BSC arm for treatment-emergent AEs (TEAEs); more specifically, drug-related grade 3–4 TEAEs, treatment-emergent serious AEs (SAEs), and grade 3–4 treatment-emergent SAEs. The most frequently reported AEs were fatigue (39.7 % AAP vs. 34.6 % PP), back pain (33.2 vs. 33.1 %), arthralgia (29.3 vs. 24.4 %), nausea (24.0 vs. 23.0 %), peripheral oedema (26.0 vs. 20.9 %), constipation (23.6 vs. 20.4 %), diarrhoea (23.4 vs. 18.1 %), and hot flush (22.7 vs. 18.3 %). AAP resulted in significantly more grade 3 or 4 increased alanine transaminase, increased aspartate aminotransferase, and dyspnoea, but less hydronephrosis.

### 3.2 Critique of Clinical Effectiveness Evidence and Interpretation

Literature are available suggesting that docetaxel might be less effective following AA [17]. Assuming that most patients will end up using docetaxel, an important question in this appraisal is whether AAP followed by docataxel is more effective than watchful waiting (BSC) followed by docetaxel. In the COU-AA-302 trial, 239 of 546 (43.8 %) AAP patients and 304 of 542 (56.1 %) PP patients received docetaxel as subsequent therapy. The results for this specific group of patients are not presented in the CS but were submitted by the company as part of the response to the clarification letter; however, as these data were provided as commercial-in-confidence, we cannot report them here.

According to the company, the Independent Data Monitoring Committee for the COU-AA-302 trial concluded on 27 February 2012 that patients in the AAP arm had a 'highly significant advantage', even though the HR for OS had not reached the stringent prespecified statistical significance level (0.0034). The committee unanimously recommended stopping the study, unblinding, and allowing crossover. The study was unblinded on 2 April 2012, and crossover from BSC to AAP occurred following unblinding (2 April 2012) for three patients by the third interim analysis (22 May 2012). Neither the second nor third interim analysis OS results met the prespecified statistical significance levels. Because crossover was now allowed, it is unlikely that the trial will ever show a significant survival benefit.

### 3.3 Cost-Effectiveness Evidence Submitted by the Company

The CS [2] included a literature search of relevant cost-effectiveness studies; however, it did not identify any studies on AAP for the treatment of adult men who were asymptomatic or mildly symptomatic after failure of ADT and in whom chemotherapy is not yet clinically indicated. Therefore, a de novo economic analysis was performed by the company.

The company presented a comparison of AAP versus BSC by means of a discrete event simulation (DES) model, tracking patients at the individual level. The model follows patients until age 100 years, which is assumed to reflect a lifetime time horizon. Patients entering the model (Fig. 1) are assigned to either the AAP or the BSC strategy. Patients who discontinue pre-docetaxel active treatment or progress are monitored in a BSC phase before starting docetaxel. After the docetaxel treatment phase, patients are monitored in a BSC phase for progression again upon which they could receive active treatment (AAP) if deemed appropriate. However, patients who had already received AAP in the first-line are not eligible for re-treatment with AAP post-docetaxel. After all treatment options had been explored and disease has progressed, patients then enter a palliative stage. Hence, the model effectively compares AAP followed by docetaxel and subsequent treatments (not including AA) with watchful waiting (including BSC) followed by docetaxel and subsequent treatments (including AA).

The model was primarily populated using the COU-AA-302 trial (third interim analyses) [13–15] and consisted of a total of 17 prediction equations for estimating time to treatment discontinuation (TTD), time to treatment start, time to death within the various treatment phases, and (disease) status of the patient at different phases. To estimate these prediction equations, study data of 902 patients were used [83 % of the intention-to-treat (ITT) population, which consisted of 1088 patients]. Various covariates were included in these prediction equations, chosen largely on the basis of statistical significance, although the ERG noted that non-significant covariates were inconsistently included in some cases. These prediction equations were combined
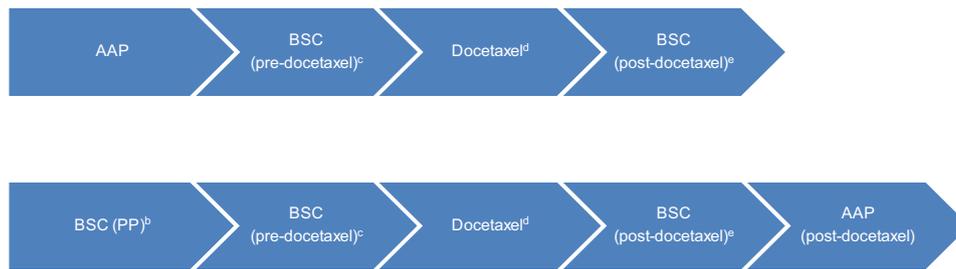
Fig. 1 Visual representation of the DES model (see Figs. 5.1 and 5.2 from the ERG report [3] for more details). *DES* discrete event simulation, *AAP* abiraterone acetate in combination with prednisone/prednisolone, *BSC* best supportive care, *PP* placebo plus prednisone/prednisolone, *ERG* Evidence Review Group, *ECOG* Eastern Cooperative Oncology Group, *PS* performance score. [a] Patients could die in all stages of the model, except during AAP, BSC (PP), and post-docetaxel treatment. If patients die, they firstly go through the 'BSC before death' phase involving palliative care, until death. This consists of the 'end-of-life' phase where patients are near death and will not receive additional active treatments that may impact survival, but instead are managed for their pain or other symptoms. [b] BSC (PP) involves active monitoring without active treatments that impact survival (patients are still receiving treatments that palliate symptoms of disease, e.g. corticosteroids). [c] Patients for whom pre-docetaxel treatment was discontinued or in whom disease was progressed were monitored in a BSC (pre-docetaxel) phase prior to commencing docetaxel treatment. No active treatment that impacted survival was provided during this phase (although patients are still receiving treatments that palliate symptoms of disease). [d] Patients started docetaxel only if ECOG PS score <2 (assumed to correspond to Karnofsky PS score ≥60 %). Otherwise, patients moved to 'BSC before death' until death. [e] This phase involves no active treatment that has shown to impact overall survival while patients are still receiving treatments that palliate symptoms of disease. Furthermore, it was assumed that if patients received AAP prior to docetaxel they would not be eligible for AAP retreatment post-docetaxel, whereas BSC patients were allowed to receive AAP post-docetaxel

with the profile/characteristics of individual patients to estimate the exact treatment path, including duration in the various treatment phases, and survival.

Although utility data were obtained from the COU-AA-302 trial (indirectly via mapping FACT–P results) [13–15], utility values in the base-case model came from a UK mCRPC utility study (online survey among 163 patients). Only the base case on-treatment utility increment of AAP over BSC (pre-docetaxel) was obtained from the COU-AA-302 trial [13–15]. For all other treatment phases, FACT–P-mapped utilities were included in a scenario analysis. AEs were not separately taken into account in the utility score as the safety profile of AAP and BSC was considered similar, and all other effects of treatment (e.g. docetaxel) on HRQoL would have been captured in the treatment-phase specific utility value. No utility increment was applied for post-docetaxel AAP treatment, unlike STA259 (considering AAP for mCRPC previously treated with docetaxel) [5]. For advanced/metastatic prostate cancer, utility values typically vary between 0.50 and 0.87 [18].

Costs (2012–2013 price level) were considered from an NHS and Personal Social Services perspective. Moreover, costs were subdivided into treatment costs, costs of scheduled medical resource utilisation (MRU), and costs of unplanned MRU (including AEs). Monthly treatment costs for AAP are considerably higher than the cost for BSC, which was represented by prednisolone 10 mg daily and is therefore negligible. The monthly cost of docetaxel, including administration costs, is £1550. Scheduled MRU was assessed by means of a survey among 53 UK oncologists, with questions on total outpatient visits, scans, and laboratory tests. For AAP-treated patients, both pre- and

post-docetaxel, a higher MRU is applied until 3 months after the start of treatment because they require additional monitoring. Unplanned events while on treatment were estimated, where possible, based on the COU-AA-301 [19, 20] and COU-AA-302 [13–15] trial data. However, since these trials did not contain unplanned MRU data for BSC (pre- and post-docetaxel and palliative phase), or docetaxel, unplanned MRU of proxy groups had to be used for these phases in the model. For pre- and post-docetaxel phases, treatment of AEs was considered to be included in the unplanned MRU. Costs of incremental grade 3 or 4 AEs for docetaxel compared with AAP were assigned separately. Resources and medication used for treating these AEs were assessed by means of expert opinion.

The base-case deterministic incremental cost-effectiveness ratio (ICER) for AAP versus BSC was £46,722 per QALY gained. One-way sensitivity analyses indicated that the most influential parameters are likely to be the post-ADT baseline utility and the discount rate for the health benefits. In addition, scenario analyses were performed on various assumptions. When excluding the PAS, and also in the scenario where FACT–P mapping utilities were used instead of EQ-5D from the patient utility study, this resulted in ICERs above £50,000 per QALY gained. For all other scenarios, ICERs would be lower than £50,000 per QALY gained.

## 3.4 Critique of Cost-Effectiveness Evidence and Interpretation

The critical appraisal of the company's economic evaluation by the ERG highlighted a number of concerns:

- deviation from the decision problem defined in the scope;
- overly complicated model that lacks transparency;
- using the analysable dataset instead of the ITT population;
- inconsistencies in the estimation of prediction equations;
- not fully incorporating the impact of AEs;
- on-treatment utility increment for post-docetaxel AA;
- short post-docetaxel survival

The main deviation from the decision problem defined in the scope [6] was that docetaxel is not included as a comparator. However, as the indication is men with mCRPC in whom chemotherapy is not yet clinically indicated, it seems reasonable that docetaxel is not considered as a comparator.

Regarding the model structure, the ERG does not believe that a DES model, simulating individual patients by means of 17 prediction equations, was the most transparent approach possible to address the decision problem defined in the scope. Transparency is a key aspect of modelling and in this specific case a more transparent model would be more convenient for an external reviewer to assess face validity and internal validity of the model. Moreover, a more transparent model would have allowed the ERG much more flexibility in performing additional analyses.

The prediction equations were estimated based on what the company referred to as the analysable patient sample, which is a subset ($n = 902$) of the ITT population ($n = 1088$). The company argued that the ITT population could not be used for estimating prediction equations because baseline data were missing for a number of patients. However, this approach introduced bias in favour of AAP for both TTD and OS (as OS is dependent on TTD). This is illustrated in Fig. 3 in the company's response to NICE's request for additional information (see Janssen [16]). Therefore, the ERG would have preferred an approach in which the prediction equations were based on the ITT population and imputing any missing baseline data or, alternatively, to use treatment as the only covariate.

In addition, the process of estimating the prediction equations was not consistent. For instance, the equation for 'time from AAP/BSC (PP) end to death' was, unlike all other prediction equations, estimated separately by arm, while for all other equations, treatment was included as a covariate. Although requested by the ERG in the clarification phase, the company could not provide a convincing reason for using this procedure [16]. Furthermore, candidate covariates varied between prediction equations. A rationale for selecting the candidate covariates was absent. In addition, interaction terms were sometimes included in an equation despite a non-significant $p$-value. Adding covariates or interaction terms even when they were not statistically significant for 'time to AAP/BSC (PP) end' and 'time from post-docetaxel treatment end to death' could not be regarded as conservative as this increased the effectiveness of AAP versus BSC in both instances (see Sect. 5.2.6 of the ERG report [3] for more details). Therefore, the ERG would have preferred a well-defined and consistently applied procedure on whether or not to stratify, and on including covariates and interaction terms. Without such a procedure, it is difficult to rule out bias caused by these elements.

Although AAP seemed to be associated with more grade 3 and 4 AEs, the company argued that, because AAP and BSC have a similar safety profile, differential AE utility values for AAP and BSC were not indicated, and the on-treatment utility gain for AAP versus BSC would capture all relevant differences. The only way AEs were explicitly taken into account was in the costs of treating AEs during the docetaxel phase. Therefore, AEs were not incorporated separately in HRQoL in anyway, nor were they incorporated in the costs in the pre- and post-docetaxel phases. In the clarification phase, the ERG requested an additional analysis, removing the on-treatment utility gain and using per AE utility decrements, as well as pre- and post-docetaxel AE treatment costs [16]. The ICER in this additional analysis increased to £50,880, indicating that not explicitly incorporating AE utility decrements is not conservative. In addition, SREs were not considered by the company, whereas they were included in STA259 [5] and mentioned in the scope [6]. Given that COU-AA-301 [19, 20] demonstrated that, for post-docetaxel AAP, time to SREs was improved compared with placebo, it can be questioned whether not including SREs in the present submission can be considered conservative.

Unlike in STA 259 [5], no post-docetaxel on-treatment utility increment for AAP was applied in the current assessment. The company argued that applying a post-docetaxel utility increment of 0.046 (derived from COU-AA-301 trial data) would be double counting since the majority of patients in the UK mCRPC utility study were assumed to have already been receiving AAP in this setting, and therefore the on-treatment utility gain was captured directly in the utility value. However, the ERG could not see any reason why this would not still allow the use of a differential utility value, and requested an analysis incorporating a BSC on-treatment decrement. The company performed this analysis, together with a higher post-docetaxel baseline utility, to be more in line with STA259 [5] (also requested by the ERG). This analysis resulted in an ICER of £47,936 per QALY gained. The ERG therefore concluded that the results are rather robust with respect to these changes in utility values post-docetaxel.

Post-docetaxel survival in the current model seems very low compared with STA259 [5]. This is difficult to explain

given that STA259 considered patients who were in the post-docetaxel phase.

## 3.5 Additional Work Undertaken by the ERG

Due to the abovementioned concerns, the ERG questioned the validity of the ICER provided by the company. The ERG was able to resolve some of the issues highlighted by using an on-treatment utility for post-docetaxel active treatment, and non-stratified prediction equations based on the ITT population, using treatment as the only covariate. This resulted in an ICER of £57,688 per QALY gained for the ERG base case (Table 1).

ICERs calculated in the additional sensitivity analyses performed by the ERG ranged between £56,671 and £74,803 per QALY gained. Assuming post-docetaxel survival is equal to that in STA259 [5] (by adjusting the coefficients for 'time from post-docetaxel treatment discontinuation to death') resulted in an ICER of £65,515 per QALY gained. Finally, replacing the log-logistic distributions (two prediction equations) with Weibull distributions resulted in an ICER of £74,803 per QALY gained (Table 1).

## 3.6 End-of-Life Criteria

NICE end-of-life supplementary advice should be applied in the following circumstances and when all the criteria referred to below are satisfied [21]:

- the treatment is indicated for patients with a short life expectancy, normally less than 24 months;
- there is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment;
- the treatment is licensed, or otherwise indicated, for small patient populations.

With regard to the first criterion, the CS [2] showed that after 24 months, approximately 63 % of subjects in the control group are still alive, and that the median survival is 30.1 months (95 % CI 27.3–34.1). Therefore, it was unlikely that life expectancy in this patient group would be less than 24 months. According to the company, patients in the trial were likely to have gone on to receive other clinical trial technologies post-docetaxel, and therefore the survival observed for these patients was probably not reflective of the average mCRPC patient in the UK. However, as far as the ERG was aware, the 'short life expectancy, normally less than 24 months' was based on the normal treatment options available for these patients without the intervention under assessment.

With regard to the second criterion, the company provided median survival estimates, but not mean survival, in the CS [2]. In the clarification letter, the ERG asked the company to provide the mean survival in the BSC group in COU-AA-302 for the overall population and for the subgroup of patients from UK centres, and the mean survival

**Table 1** Overview of additional analyses undertaken by the ERG (reported in original ERG report)

| Technology | Incremental costs (£) | Incremental QALYs | ICER, £/QALY gained |
|---|---|---|---|
| Company base case | 26,404 | 0.57 | 46,722 |
| Recalculated company base case[a] | 26,423 | 0.57 | 46,756 |
| Post-docetaxel on treatment utility[b] | 26,423 | 0.56 | 46,952 |
| Updated prediction equations[c] | 24,757 | 0.43 | 57,337 |
| ERG base case[d] | 24,757 | 0.43 | 57,688 |
| Additional sensitivity analyses (based on ERG base case) | | | |
|    Remove cabazitaxel negative treatment effect | 24,821 | 0.44 | 56,671 |
|    Equal post-docetaxel survival compared with STA259 | 24,159 | 0.37 | 65,515 |
|    Weibull instead of log-logistic | 19,620 | 0.26 | 74,803 |
|    Weibull instead of log-normal | 24,565 | 0.43 | 57,202 |

*QALY* quality-adjusted life-year, *ICER* incremental cost-effectiveness ratio, *ERG* Evidence Review Group, *STA* NICE single technology appraisal, *AA* abiraterone acetate, *BSC* best supportive care, *CS* company submission

[a] The ERG could not replicate the results presented in the CS using the economic model provided by the company. This minor deviation was later explained by a typographical mistake prior to submission by the company, resulting in higher docetaxel adverse event costs being double counted when the ERG ran the model

[b] A disutility of 0.046 was applied in the post-docetaxel phase for patients not receiving active treatment (i.e. receiving BSC instead of AA)

[c] Prediction equations based on the intention-to-treat population, and including treatment as the only covariate, were used (based on the '302 mode Parametric Functions Parameters' file provided by the company in response to clarification question B4a)

[d] A combination of the two scenarios mentioned above

gain of AAP compared with BSC in COU-AA-302 for the overall population and for the subgroup of patients from UK centres. The company responded that they were unable to answer these questions (see company response to request for clarification from the ERG [16]).

With regard to the third criterion, it is likely that the treatment is indicated for a small patient population.

### 3.7 Conclusion of the ERG Report

An important question in this appraisal was, according to the ERG, whether AAP followed by docetaxel is more effective than BSC followed by docetaxel. In the COU-AA-302 trial, 239 of 546 (43.8 %) AAP patients and 304 of 542 (56.1 %) PP patients received docetaxel as subsequent therapy, following AA or placebo [2]. The results for this specific group of patients were not presented in the CS; therefore, the ERG asked the company to provide these data in the clarification letter. However, these data were presented as commercial-in-confidence and can therefore not be reported here.

The ERG was able to resolve some of the issues highlighted in the cost-effectiveness section of the report, and calculated an ICER of £57,688 per QALY gained for the ERG base case. This included using the ITT population, with treatment as the only covariate, for the estimation of the prediction equations. Ideally, the ERG would have preferred an approach in which the prediction equations are based on the ITT population and imputing any missing baseline data to be able to consistently use additional covariates. However, the ERG was unable to provide these analyses as it did not have access to the individual patient-level data. Moreover, the ERG acknowledged that uncertainties remain concerning the reliability of the cost-effectiveness evidence, which could neither be handled in the ERG base case nor could a sensitivity analysis be provided to estimate the impact of these issues on the results. These issues include not including the possibility of dying during AAP/BSC treatment and post-docetaxel active treatment, not using differential costs and utilities for all AEs for all treatment phases (including SREs), and lack of empirical data to calculate resources and costs for most of the treatment phases.

### 3.8 Addenda Submitted by the ERG

Not including separate additional analyses requested by NICE, the ERG submitted four addenda [4] (upon request), in response to questions raised by the AC, to additional data provided by the company and to a new PAS submitted by the company. Moreover, this appraisal entailed, in total, five AC meetings (ACMs), two ACDs, and two FADs. For clarity, a timeline is provided in Table 2.

In its response to the second ACD, the company provided new cost-effectiveness analyses, including the company's base case, resulting in an ICER of £28,563 per QALY gained. The following changes were applied to the company's original base case to obtain this new base case:

- a new PAS was incorporated;
- the docetaxel drug price was reduced by 20 %;
- a utility increment of 0.021 was applied to the post-docetaxel active treatment phase of the model.

The new PAS included a permanent reduction in the official list price for AA by 21.5 %, resulting in a price of £2300 per 30-days pack [22]. In addition, as part of the new PAS, the drug acquisition costs of AA would be rebated to the NHS after 10 months of treatment for each individual patient.

In addition to the new base case, the company included a piecewise curve (log-logistic + Weibull for extrapolation [4]) for time to first-line TTD instead of using the log-logistic distribution only. This was done since it is unclear whether the long tail of the log-logistic distribution is clinically plausible. This analysis also included arbitrarily limiting TTD to a maximum of 1000 days for BSC only, and resulted in an ICER of £32,849 per QALY gained.

The ERG base case included some additional adjustments to the original company base case:

- Including a PAS administration fee as part of the new PAS.
- Using the old PAS for BSC (post-docetaxel AA) and the new PAS for AA (pre-docetaxel AA). This was preferred because, in case the current appraisal does not recommend AA before docetaxel, the old PAS would be maintained for AA after docetaxel.
- Assuming that AA non-compliance does not lead to recoverable drug costs.
- Applying a utility increment of 0.046 (consistent with STA259 [5]) to the post-docetaxel active treatment phase of the model.
- Using the prediction equations based on the ITT population, including treatment as the only covariate.

Using these additional adjustments resulted in an ICER of £38,061 per QALY gained for the ERG base case and £54,091 per QALY gained when also using the piecewise curve for TTD (without the arbitrary TTD cap of 1000 days for BSC) [see Table 3 for an overview of selected analyses based on the new PAS].

## 4 Key Methodological Issues

The ERG raised several issues regarding the cost-effectiveness analyses methods and assumptions presented by the company. The impact of some of these issues on the

**Table 2** Timeline

| Date | Event | Comment |
|---|---|---|
| December 2012 | Marketing authorisation AA extended to pre-docetaxel | |
| April 2014 | ACM 1 | ACD 1: not recommended (AC concluded that all the ICERs estimated by both the company and the ERG fell substantially above the range normally considered cost effective, i.e. £20,000–30,000 per QALY gained) |
| June 2014 | ERG addendum 1 | ERG response to comments raised during ACD consultation on the benefit of delaying chemotherapy |
| June 2014 | ACM 2 | FAD 1: not recommended (AC concluded that all the ICERs estimated by both the company and the ERG fell substantially above the range normally considered cost effective, i.e. £20,000–30,000 per QALY gained) |
| September 2014 | Appraisal suspended pending revised PAS | |
| April 2015 | ERG addendum 2 | ERG response to new PAS |
| October 2015 | ACM 3 | Request new data from company |
| November 2015 | ERG addendum 3 | ERG response to the company's response to NICE's request for additional information |
| November 2015 | ACM 4 | ACD 2: not recommended (AC most plausible ICER likely between £35,500 and £59,600 per QALY gained) |
| January 2016 | ERG addendum 4 | ERG response to the company's response to ACD |
| February 2016 | ACM 5 | FAD 2: recommended (AC most plausible ICER likely between £28,600 and £32,800 per QALY gained; see Sect. 5) |

*AA* abiraterone acetate, *ACM* Appraisal Committee Meeting, *ACD* Appraisal Consultation Document, *AC* Appraisal Committee, *ICER* incremental cost-effectiveness ratio, *ERG* Evidence Review Group, *QALY* quality-adjusted life-year, *PAS* patient access scheme, *FAD* Final Appraisal Determination, *NICE* National Institute for Health and Care Excellence

**Table 3** Selected additional analyses undertaken by the company and the ERG (using the new PAS)

| Technology | Incremental costs (£) | Incremental QALYs | ICER, £/QALY gained |
|---|---|---|---|
| Company base case | 16,055 | 0.56 | 28,563 |
| Company base case + piecewise curve | 15,855 | 0.48 | 32,849 |
| ERG base case[a] | 15,089 | 0.43 | 35,486 |
| ERG base case[b] | 16,098 | 0.43 | 37,859 |
| ERG base case[c] | 16,184 | 0.43 | 38,061 |
| ERG base case + piecewise curve | 15,938 | 0.31 | 51,026 |
| ERG base case + piecewise curve (without arbitrarily limiting TTD to a maximum of 1000 days for BSC only) | 15,908 | 0.29 | 54,091 |

*PAS* patient access scheme, *QALY* quality-adjusted life-year, *ICER* incremental cost-effectiveness ratio, *ERG* Evidence Review Group, *TTD* time to treatment discontinuation, *BSC* best supportive care, *AA* abiraterone acetate, *AAP* AA in combination with prednisone/prednisolone

[a] ERG base case, including the new PAS for both BSC (post-docetaxel AA) and AAP (pre-docetaxel AA), not including the PAS administration fee (for the new PAS)

[b] ERG base case, including the old PAS for BSC (post-docetaxel AA) and the new PAS for AA (pre-docetaxel AA), not including the PAS administration fee (for the new PAS)

[c] ERG base case, including the old PAS for BSC (post-docetaxel AA) and the new PAS for AA (pre-docetaxel AA), including the PAS administration fee (for the new PAS)

estimated ICER was examined by the ERG and, if applicable, adjusted in the ERG base case. The issue that appeared to have the most impact on the ICER was using the ITT population and the consistent selection of (candidate) covariates and interaction terms for estimating the prediction equations. This was salvaged in the ERG base

case by using the ITT population to estimate the prediction equations, and including treatment as the only covariate. Ideally, the approach to estimate the prediction equations would be based on the ITT population and imputing any missing baseline data to be able to consistently select covariates and/or interaction terms.

The ERG acknowledged that there are remaining uncertainties that could not be examined and/or included in the ERG base case, including censoring for BSC patients after sequential treatment with AAP and cabazitaxel, not including the possibility of dying during AAP/BSC treatment and post-docetaxel active treatment, not using differential costs and utilities for all AEs for all treatment phases, and no empirical data to calculate resources and costs for most of the treatment phases. Moreover, during the ACMs, the face validity of the economic model was questioned because clinical experts stated that patients switch from first-line treatment to docetaxel within 1 week of disease progression, whereas this was estimated to be over 5 months in the model (see second ACD and FAD [4]).

## 5 National Institute for Health and Care Excellence (NICE) Guidance

In March 2016, the AC produced the final guidance, stating that AAP is recommended, within its marketing authorisation, as an option for treating mCRPC:

- In people who have no or mild symptoms after ADT has failed, and before chemotherapy is indicated.
- Only when the company rebates the drug cost of AA from the eleventh month until the end of treatment for people who remain on treatment for more than 10 months.

### 5.1 Consideration of Clinical Effectiveness

The AC concluded that AAP delayed disease progression and improved OS compared with placebo, but that there was uncertainty about the extent of the survival benefit. The AC stated that as chemotherapy can reduce a person's quality of life, treatments delaying the need for chemotherapy are highly valued by patients.

The AC noted that AAP was innovative and that the utility values in the model may not fully capture the benefit to patients of delaying cytotoxic chemotherapy.

The AC concluded that current mean life expectancy for chemotherapy-naïve mCRPC was unlikely to be less than 24 months, and AAP at this stage in the treatment pathway did not meet the end-of-life criterion for short life expectancy.

### 5.2 Consideration of Cost Effectiveness

The AC noted that the scope (issued by NICE in 2012) included docetaxel as a comparator, but the company did not include docetaxel as a comparator because the marketing authorisation states that AA should be used for people for whom chemotherapy is not yet indicated. The AC agreed that not incorporating docetaxel was appropriate.

The AC stated that a DES model was not unreasonable, but that the company's model was particularly complex and lacked transparency, which made it difficult for the ERG to validate and critique, and for the AC to determine the plausibility of the model outcomes.

In principle, the AC agreed with the ERG that it is preferable to use the ITT population for modelling because this reduces the risk of bias. However, in this specific case, the AC agreed with the company that it was appropriate to use the full covariate subgroup rather than the ITT population.

With regard to the 17 equations predicting time to events or disease status in the DES model, the AC noted that the company made a large number of judgements when determining which covariates to include and which parametric distribution to choose for extrapolation. The AC noted that, for two equations, the company had not followed its own statistical plan when choosing covariates, and the AC agreed with the ERG that this could introduce bias to the model.

For estimation of TTD for the duration of the trial period, the AC stated that the log-logistic curve was the best fit to the trial data (used in the base case of both the company and the ERG). However, it was noted that the log-logistic curve predicted that some patients remained on AAP for a long time (4 % took AAP for 8 years or more), which could not be supported by trial data. The AC noted that the Weibull curve predicted that fewer patients remain on AAP in the long term, and the piecewise curve gave predictions that were in-between the log-logistic and Weibull curves. However, in this latter analysis, the company assumed that all patients stopped having BSC at 1000 days, and the AC was concerned that this assumption may not be clinically plausible. The AC concluded that, for predicting TTD, it is preferable to use either the log-logistic or piecewise curve.

The AC discussed post-docetaxel survival estimates of the DES model. It was noted that the company had not used data from TA259 [5] to check the validity of its model in the current appraisal. It also noted that the modelled post-docetaxel survival times were shorter in the current appraisal (based on data from COU-AA-302 [13–15]) than in TA259 [5] (based on data from COU-AA-301 [19, 20]). The AC was aware that the ERG carried out a scenario

analysis in which it fixed post-docetaxel survival to be the same as in COU-AA-301, and this subsequently increased the ICER. The AC heard from the company that, although the estimates from COU-AA-301 came from a larger sample of patients, it did not consider these data to be relevant for the current appraisal because the population in COU-AA-301 was different from that in COU-AA-302. On balance, the AC concluded that it was appropriate to use COU-AA-302 to estimate post-docetaxel survival times. Nonetheless, it was also concluded that uncertainty about the modelled survival times persisted because only a small number of patients from COU-AA-302 contributed data to this phase of the model.

The AC understood that the company's base-case model used 98 % of the cost of the licensed dose of AA, as in COU-AA-302 [13–15] patients took, on average, 98 % of the licensed dose. The AC considered that the full cost of the licensed dose of AA should be included in the model as the cost of unused tablets was unlikely to be recovered. Additionally, it was noted by the AC that the costs of administering the PAS, although low, had not been included in the company's base case, and considered that these costs should have been included. Moreover, the AC acknowledged that the two PASs would not and could not exist at the same time. Nonetheless, it concluded that it was appropriate to include the existing PAS for BSC and the new complex PAS for AA for the purposes of decision making, and it acknowledged that using this approach in a scenario analysis had a modest impact on the ICER. The AC noted that the company's assumptions relating to these costs favoured AA, but that including the AC preferred assumptions increased the ICER for AA, P compared with BSC, only slightly.

## 6 Conclusions

This paper describes the STA on AAP for the treatment of chemotherapy-naïve mCRPC. The evidence suggests that AAP is an effective treatment option for the treatment of chemotherapy-naïve mCRPC. The preferred analysis of the ERG showed that AAP might not be cost effective, but the AC did not agree with the ERG assumption to use the ITT population and considered that the most plausible ICERs were within the range that could be considered cost effective, and hence recommended AAP for chemotherapy-naïve mCRPC. However, it should be noted that this STA did not consider enzalutamide, which is now recommended as an option for treating chemotherapy-naïve mCRPC, and hence no statements can be made regarding the cost effectiveness of AAP versus enzalutamide for this population.

## References

1. National Institute for Health and Care Excellence (NICE). Guide to the processes of technology appraisal. London: NICE; 2014. https://www.nice.org.uk/article/pmg19/chapter/foreword. Accessed 21 Mar 2016.
2. Janssen. Abiraterone acetate (Zytiga®) for the treatment of metastatic castration-resistant prostate cancer in men not previously treated with chemotherapy: submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA). Janssen; 2014.
3. Riemsma R, Ramaekers BLT, Tomini F, Wolff R, van Asselt ADI, Joore MA, et al. Abiraterone for the treatment of chemotherapy naïve metastatic castration-resistant prostate cancer: a single technology appraisal. York: Kleijnen Systematic Reviews Ltd; 2014.
4. National Institute for Health and Care Excellence (NICE). Abiraterone for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated. London: NICE; 2016. https://www.nice.org.uk/guidance/TA387. Accessed 21 Mar 2016.
5. National Institute for Health and Care Excellence (NICE). Abiraterone for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen. London: NICE; 2012. https://www.nice.org.uk/guidance/ta259/history. Accessed 21 Mar 2016.
6. National Institute for Health and Care Excellence (NICE). Abiraterone acetate for the treatment of metastatic hormone relapsed prostate cancer not previously treated with chemotherapy: single technology appraisal. Final scope. London: NICE; 2013. https://www.nice.org.uk/guidance/GID-TAG434/documents/prostate-cancer-metastatic-castrateresistant-not-treated-with-chemotherapy-

abiraterone-acetate-with-prednisolone-appendix-b-final-scope2. Accessed 21 Mar 2016.

7. National Institute for Health and Care Excellence (NICE). Prostate cancer: diagnosis and treatment, NICE clinical guideline 175. London: NICE; 2014. https://www.nice.org.uk/guidance/cg175. Accessed 21 Mar 2016.

8. Vasaitis TS, Bruno RD, Njar VC. CYP17 inhibitors for prostate cancer therapy. J Steroid Biochem Mol Biol. 2011;125(1–2):23–31.

9. eMedicine from WebMD. Prostate cancer diagnosis and staging. Medscape: eMedicine. 2016. http://emedicine.medscape.com/article/458011-overview#showall. Accessed 21 Mar 2016.

10. Rentsch CA, Cecchini MG, Thalmann GN. Loss of inhibition over master pathways of bone mass regulation results in osteosclerotic bone metastases in prostate cancer. Swiss Med Wkly. 2009;139(15–16):220–5.

11. European Medicines Agency (EMA). Zytiga. Annex 1: summary of product characteristics. London: EMA; 2014. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002321/WC500112858.pdf. Accessed 21 Mar 2016.

12. National Institute for Health and Care Excellence (NICE). Enzalutamide for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated. London: NICE; 2016. https://www.nice.org.uk/guidance/ta377/history. Accessed 21 Mar 2016.

13. Ryan CJ, Smith MR, Fizazi K, Saad F, Mulders PFA, Sternberg CN, et al. Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naive men with metastatic castration-resistant prostate cancer (COU-AA-302): final overall survival analysis of a randomised, double-blind, placebo-controlled phase 3 study. Lancet Oncol. 2015;16(2):152–60.

14. Rathkopf DE, Smith MR, de Bono JS, Logothetis CJ, Shore ND, de Souza P, et al. Updated interim efficacy analysis and long-term safety of abiraterone acetate in metastatic castration-resistant prostate cancer patients without prior chemotherapy (COU-AA-302). Eur Urol. 2014;66(5):815–25.

15. Ryan CJ, Smith MR, de Bono JS, Molina A, Logothetis CJ, de Souza P, et al. Abiraterone in metastatic prostate cancer without previous chemotherapy. N Engl J Med. 2013;368(2):138–48.

16. Janssen. Company response to the request for additional information. 2015. https://www.nice.org.uk/guidance/TA387/documents/appraisal-consultation-document-2. Accessed 21 Mar 2016.

17. Mezynski J, Pezaro C, Bianchini D, Zivi A, Sandhu S, Thompson E, et al. Antitumour activity of docetaxel following treatment with the CYP17A1 inhibitor abiraterone: clinical evidence for cross-resistance? Ann Oncol. 2012;23(11):2943–7.

18. Torvinen S, Bergius S, Roine R, Lodenius L, Sintonen H, Taari K. Use of patient assessed health-related quality of life instruments in prostate cancer research: a systematic review of the literature 2002–15. Int J Technol Assess Health Care. (Epub 22 Mar 2016).

19. Fizazi K, Scher HI, Molina A, Logothetis CJ, Chi KN, Jones RJ, et al. Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study. Lancet Oncol. 2012;13(10):983–92.

20. 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(21):1995–2005.

21. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. London: NICE; 2013. https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed 4 Apr 2016.

22. National Institute for Health and Care Excellence (NICE). NICE recommends abiraterone for prostate cancer. London: NICE; 2016. https://www.nice.org.uk/news/press-and-media/nice-recommends-abiraterone-for-prostate-cancer. Accessed 9 May 2016.