External Assessment Group Report commissioned by the NIHR Evidence Synthesis Programme on behalf of NICE

Darolutamide with androgen deprivation therapy for treating hormone sensitive metastatic prostate cancer

---

Produced by	**Southampton Health Technology Assessments Centre (SHTAC)**
Authors	Lois Woods, Senior Research Assistant, Evidence Synthesis and Information Specialist Marcia Takahashi, Research Fellow, Health Economics Neelam Kalita, Senior Research Fellow, Health Economics Jonathan Shepherd, Principal Research Fellow, Evidence Synthesis
Correspondence to	Dr Jonathan Shepherd Southampton Health Technology Assessments Centre (SHTAC) School of Healthcare Enterprise and Innovation University of Southampton Alpha House, Enterprise Road, University of Southampton Science Park, Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	9th June 2025 (post factual accuracy check version 26th June 2025)
Source of Funding	This report was commissioned by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme as project number NIHR175493.

Acknowledgements

We thank the following for providing expert clinical advice to the project team and for reading and commenting on a draft of the report:

Dr Mohini Varughese, Consultant Clinical Oncologist and Honorary Clinical Associate Professor, Department of Oncology, Royal Devon University Healthcare NHS Foundation Trust & Honorary Clinical Associate Professor, University of Exeter

We also thank Geoff Frampton, Principal Research Fellow, Southampton Health Technology Assessments Centre (SHTAC), and Fay Chinnery, Research Fellow, Southampton Health Technology Assessments Centre (SHTAC), for providing a quality assurance review of the draft report.

Declared competing interests of the authors and advisors

The EAG and the clinical advisor declare no competing interests.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Woods, L; Takahashi, M; Kalita, N; and Shepherd, J. Darolutamide with androgen deprivation therapy for treating hormone sensitive metastatic prostate cancer. A Cost-comparison Technology Appraisal. Southampton Health Technology Assessments Centre, 2025.

Contributions of authors

Lois Woods critically appraised the clinical effectiveness systematic review and drafted the report; Marcia Takahashi critically appraised the cost comparison model and drafted the report; Neelam Kalita critically appraised the cost comparison model and drafted the report; Jonathan Shepherd critically appraised the clinical effectiveness systematic review, drafted the report and is the project coordinator.

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ADL	Activities of daily living
ADT	Androgen deprivation therapy
AE	Adverse event
AIC	Academic in confidence
AIC	Akaike information criterion
ALT	Alanine aminotransferase
AO	Adverse outcome
APCCC	Advanced Prostate Cancer Consensus Conference
AR	Androgen receptor
ARIs	Androgen receptor inhibitors
ARTA	Androgen receptor-targeted agents
ASI	Androgen synthesis inhibitor
AST	Aspartate aminotransferase
BBB	Blood-brain barrier
BICR	Blinded independent central review
BNF	British National Formulary
BPI-SF	Brief pain inventory – short form
BRMA	Bayesian bivariate random-effects meta-analysis
BSC	Best supportive care
CI	Confidence interval
CIC	Commercial in confidence
CNS	Central nervous system
CPI	Consumer price inflation
CR	Complete response
CrI	Credible interval
CRD	Centre for Reviews and Dissemination
CRPC	Castration resistant prostate cancer
CS	Company submission
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CVD	Cardiovascular disease
CYP	Cytochrome P450
DB	Double blind
DCO	Data cut-off

DDI	Drug-drug interaction
DIC	Deviance information criteria
DOAC	Direct oral anticoagulant
DSU	Decision Support Unit
EAG	External Assessment Group
EAIR	Exposure-adjusted incidence rate
ECOG	Eastern Cooperative Oncology Group
ECOG-PS	Eastern Cooperative Oncology Group Performance Score
eGFR	Glomerular filtration rate
EMC	Electronic Medicines Compendium
EPAR	European Public Assessment Report
EQ-5D-3L	European Quality of Life Working Group Health Status Measure 3 Dimensions, 3 Levels
EQ-5D-5L	European Quality of Life Working Group Health Status Measure 5 Dimensions, 5 Levels
EQ-VAS	EuroQol Visual Analogue Scale
ESMO	European Society for Medical Oncology
FAS	Full analysis set
FACT-P	Functional Assessment of Cancer Therapy – Prostate
GABAA	γ-aminobutyric acid type A
GnRH	Gonadotropin-releasing hormone
GP	General practitioner
HCP	Healthcare provider
HR	Hazard ratio
HRCU	Healthcare resource use
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IPCW	Inverse probability of censoring weights
IPD	Individual patient level data
IPE	Iterative parameter estimate
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention to treat

IWRS	Interactive web response system
KM	Kaplan Meier
LHRH	Luteinising hormone-releasing hormone
LYG	Life-year gained
mCRPC	Metastatic castration-resistant prostate cancer
MedDRA	Medical dictionary for regulatory activities
MHRA	Medicines and Healthcare Products Regulatory Agency
mHSPC	Metastatic hormone-sensitive prostate cancer
mITT	Modified intent to treat
mpMRI	Multiparametric magnetic resonance imaging
MRI	Magnetic resonance imaging
NA	Not applicable
NCI	National Cancer Institute
NCI-CTCAE	National cancer institute common terminology criteria for adverse events
NPCA	National Prostate Cancer Audit
NE	Not estimated
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
nmCRPC	Non-metastatic castration-resistant prostate cancer
nmHSPC	Non-metastatic hormone-sensitive prostate cancer
NMA	Network metro analysis
NMB	Net monetary benefit
NCI-CTCAE (v. 5.0)	National Cancer Institute– Common Terminology Criteria for Adverse Events (version 5.0)
NR	Not reported
OATP	Organic anion-transporting polypeptide
OL	Open-label
OLE	Open-label extension
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PC	Prostate cancer
PCWG3	Prostate cancer working group 3
PD	Progressive disease

PFS	Progression-free survival
PFS2	Second progression-free survival
P-gp	P-glycoprotein 1
PRO	Patient reported outcome
PSA	Prostate-specific antigen
PSS	Personal social services
PSSRU	Personal social services research unit
PrSA	Probabilistic sensitivity analysis
PY	Patient years
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
RoW	Rest of the world
rPFS	Radiographic/radiological progression-free survival
RPSFT	Rank preserving structural failure time
RR	Relative risk/risk ratio
SAF	Safety analysis set
SAP	Statistical analysis plan
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristics
SRE	Skeletal-related events
SSE	Symptomatic skeletal event
SUCRA	Surface under the cumulative ranking curve
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
ToT	Time on treatment
TSD	Technical Support Document
UK	United Kingdom
US	United States
VAT	Value tax added
WTP	Willingness to pay

1 EXECUTIVE SUMMARY

1.1 Summary of the EAG's view of the company's cost-comparison case

Table 1 provides the EAG's bottom line view regarding the validity of the company's case for cost comparison. As can be seen, the EAG considers the criteria have been met, notwithstanding a degree of uncertainty relating to the evidence for similarity of darolutamide to the chosen comparator treatment, apalutamide (this is explained in more detail in this report).

Table 1 Criteria for cost-comparison technology appraisal

Criteria	Criteria met?	EAG considerations
The technology's expected licensed indication is the same as the chosen comparators	Yes	Darolutamide anticipated marketing authorisation is identical to that of the chosen comparator treatment, apalutamide + ADT. Specifically, darolutamide is intended for 'adult men for the treatment of metastatic hormone-sensitive prostate cancer (mHSPC) in combination with androgen deprivation therapy'.
The chosen comparators meet NICE's criteria for cost-comparison	Yes	Doublet therapy with an androgen receptor-targeted agent (ARTA) in combination with androgen deprivation therapy (ADT) is now considered first line standard care in most patients with mHSPC. The two current NICE recommended doublet therapies are apalutamide + ADT and enzalutamide + ADT. The company have chosen apalutamide as their cost comparator, providing an acceptable justification. However, they give little consideration of enzalutamide, and the reasons for or against its potential inclusion as a comparator. Expert clinical opinion to the EAG suggests that both apalutamide and enzalutamide are

Criteria	Criteria met?	EAG considerations
		commonly used as doublet therapy with ADT in practice. The relative market share of the treatments is currently unknown. The EAG considers both doublet therapies are appropriate for cost-comparison. The company is permitted to select just one comparator or more than one comparator.
It is plausible that the technology may incur similar or lower costs compared with the comparators.	Yes	Darolutamide appears to have a better adverse event profile compared to apalutamide, as suggested by the company's indirect treatment comparison (ITC). This is likely to result in reduced resource use and costs for treatment and monitoring. However, we are unable to test this assumption in relation to other cost and resource parameters as the company's economic model is not structured accordingly.

1.2 The decision problem: summary of the EAG's critique

The company's decision problem adheres to the NICE scope, with a couple of exceptions.

Firstly, the population in the NICE scope is 'People with hormone-sensitive metastatic prostate cancer', whereas the decision problem population is 'adult men with mHSPC who are unsuitable for chemotherapy'. The NICE scope does, however, state the relevant comparators as being 'For people in whom docetaxel is not suitable'. The company provide a rationale for darolutamide + ADT positioned as a treatment option in patients who are unsuitable for docetaxel, including the fact that this aligns with NICE guidance for apalutamide + ADT, their chosen cost comparator. The EAG considers the company's proposal for darolutamide in docetaxel ineligible patients is reasonable, though the anticipated marketing authorisation does not restrict the use of darolutamide to a docetaxel ineligible population. The EAG notes that the clinical effectiveness evidence for darolutamide, the ARANOTE trial, did not explicitly define the participants as being docetaxel eligible/ineligible. Rather, the trial appears to include an all-comer population.

Secondly, the decision problem does not include the two subgroups of interest listed in the scope. That is, people with newly diagnosed metastatic prostate cancer and people with high-risk metastatic prostate cancer. The company notes that most patients in the ARANOTE trial (72.5%) had newly identified mHSPC and the clinical outcomes for this subgroup are consistent with the whole trial population. Hence, a subgroup analysis would add little new information. The company also discuss the challenges in defining high risk disease and note the absence of high-risk patients in the ARANOTE trial. Expert clinical advice to the EAG suggests there is variability in practice in how high-risk patients are identified. The EAG agrees with the company's decision to not include the subgroups.

1.3 The clinical effectiveness evidence: summary of the EAG's critique

The company's pivotal phase III trial, ARANOTE, is a multi-centre, double blind study with relevant outcome measures including radiological progression free survival (rPFS) and overall survival (OS). Despite there being no UK participants recruited, the trial can be considered generally representative of the mHSPC population seen in routine NHS practice. The trial did not recruit docetaxel-ineligible patients though prior prostate cancer treatment with docetaxel or immunotherapy was not permitted.

The ARANOTE trial demonstrated statistically superior efficacy of darolutamide + ADT compared to placebo + ADT at the primary completion analysis, triggering unblinding of the trial and patient crossover from the placebo + ADT arm into the darolutamide + ADT arm. The effect of crossover is potential confounding of the differences in OS between the trial arms at the final analysis. The overall survival (OS) analyses were adjusted for crossover in sensitivity analyses and the results were consistent with the ITT analysis. However, the OS data remains immature with few events and must be interpreted with caution.

The company conducted an indirect treatment comparison (ITC) to compare the relative efficacy of darolutamide + ADT versus apalutamide + ADT. Standard methods were used to construct the NMAs and were clearly reported. The ITC results showed a trend towards favouring apalutamide + ADT for rPFS and OS, and a trend favouring darolutamide for quality of life and adverse events. Only in the comp any base case did results show any statistically significant differences: for time to deterioration in the Functional Assessment of Cancer Therapy – Prostate (FACT-P) score and for discontinuation due to adverse events; but as these differences are in favour of darolutamide + ADT a cost-comparison analysis remains appropriate.

1.4 The cost-comparison evidence: summary of the EAG's critique

The company provided a cost comparison model that estimated only the difference in the drug acquisition costs between the darolutamide + ADT and apalutamide + ADT. The EAG were unable to test the impact of varying the parameter inputs, such as resource use, subsequent treatments, and adverse events on the overall cost-comparison results, as they were not included in the model and the CS. However, based on the clinical evidence and the clinical advice to the EAG, darolutamide + ADT is likely to have similar efficacy, similar use of resources and costs (see sections 5.1 and 5.2) to apalutamide + ADT. Therefore, we do not expect that the inclusion of these parameters would impact the results negatively.

The company's base case results suggest that darolutamide + ADT is associated [redacted] relative to apalutamide + ADT, with an incremental cost of [redacted] The EAG included a half-cycle correction to the company's model (see section 5.3 and 6.3), and the corrected results slightly decreased to [redacted] (see Table 2 below). A PAS discount for darolutamide is applied, and apalutamide and ADTs (leuprorelin, goserelin, and triptorelin) are costed at list prices. Results with price discounts for apalutamide and ADTs are reported in a separate confidential addendum.

Table 2 EAG correction to the company base case: PAS price for darolutamide and list price for apalutamide and ADT medications

	Darolutamide + ADT	Apalutamide + ADT	Difference
Company base case	[redacted]	£146,218	[redacted]
+ half-cycle correction	[redacted]	£145,022	[redacted]
EAG correction to the company base case	[redacted]	£145,022	[redacted]

Source: EAG corrected cost comparison model
PAS, Patient access scheme; ADT, Androgen deprivation therapy

We performed a range of EAG exploratory scenarios and presented them with the company scenarios in section 6.4 Table 10 using the EAG corrected cost comparison model. The scenarios that have the most significant effect on the incremental cost results are:

- Changing the ToT distribution curve distributions varied the incremental total cost from [redacted] (base case: log-logistic) to [redacted] (Gompertz distribution).
- Considering rPFS as the drug cost adjustment-based curve and testing different distributions varied the incremental total cost from [redacted] (generalised gamma distribution) to [redacted] (Gompertz distribution).

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report is a critique of the company's submission (CS) to NICE from Bayer on darolutamide with androgen deprivation therapy (ADT) for treating hormone sensitive metastatic prostate cancer (mHSPC). It identifies the strengths and weakness of the CS. A clinical expert was consulted to advise the external assessment group (EAG) and to help inform this report.

Clarification on some aspects of the CS was requested from the company by the EAG via NICE on 13th May 2025. A response from the company via NICE was received by the EAG on 28th May 2025 and this can be seen in the NICE committee papers for this appraisal.

2.2 Background

The company proposes darolutamide in combination with androgen deprivation therapy as a treatment option for people with metastatic hormone sensitive prostate cancer (mHSPC) who are unsuitable to receive docetaxel. The company considers darolutamide is more tolerable and associated with fewer adverse effects compared to current NICE recommended standard of care, apalutamide + ADT, and enzalutamide + ADT. For this reason, they consider a cost- comparison technology appraisal to be appropriate. In the following sub-sections the EAG summarises and critiques the background information on this topic presented in the company submission (CS).

2.2.1 Background information on hormone sensitive metastatic prostate cancer.

The CS gives a detailed description of prostate cancer, in terms of its incidence and prevalence, risk factors, natural history, symptoms, prognosis and socio-economic consequences. The CS notes key risk factors for prostate cancer including age, prostate-specific antigen (PSA) level, obesity, a family history of prostate cancer, and ethnicity. Notably, Black African males are at significantly higher risk than White or Asian males.

2.2.2 Background information on darolutamide with androgen deprivation therapy

Darolutamide is a non-steroidal androgen receptor inhibitor for the treatment of prostate cancer. It belongs to a group of drugs known as ARTAs (androgen receptor targeted agents). First generation ARTAs include bicalutamide, flutamide, and nilutamide. These drugs work by competitively blocking the androgen receptor. The second generation ARTAs, such as abiraterone, enzalutamide and apalutamide, differ by inhibiting the androgen receptor, preventing it from binding to androgens and promoting cancer cell growth.

The CS notes that although darolutamide has the same mechanism of action as other second generation ARTAs it is a "more polar molecule, with a flexible structure, and hydrogen bond-forming potential" (CS page 12). The CS explains that "the distinct chemical structure of darolutamide differentiates it from apalutamide and enzalutamide, and results in reduced blood-brain barrier penetration and low central nervous system side effects" (CS page 26). The EAG's clinical expert commented that currently available ARTAs such as apalutamide and enzalutamide are chemically very similar, and that darolutamide shares some similarities but structurally is slightly different with the additional advantage of low penetration of the blood brain barrier. The expert echoed the company's assertion of fewer central nervous system adverse events from darolutamide, though the expert also pointed out that some events, such as seizures, are uncommon. For example, they estimated that only one of their patients has reported a seizure in the last five years.

Darolutamide is currently recommended by NICE for two prostate cancer indications, in two prostate cancer sub-populations respectively:

- **Darolutamide in combination with ADT** for treating hormone-relapsed prostate cancer in adults at high risk of developing metastatic disease (NICE TA660). Hormone-relapsed prostate cancer is also referred to as hormone-resistant or castration-resistant cancer and occurs when the patient loses hormone sensitivity. They no longer respond to ADT and their cancer progresses further.
- **Darolutamide in combination with docetaxel and ADT** as a treatment for mHSPC (NICE TA903). This is the same population group included in the NICE scope for this current NICE appraisal. These patients are still responsive to hormone therapy but have already progressed to the metastatic stage of the disease, with some patients presenting with *de novo* metastases.

2.2.3 The position of darolutamide with androgen deprivation therapy in the treatment pathway

The CS describes the current care pathway with reference to clinical guidelines from NICE, European Society of Medical Oncology (ESMO) and the European Association of Urology, plus advice from the company's expert advisory board. In terms of initial treatment for mHSPC the CS notes the following:

- **Current ESMO guidelines recommend triplet or doublet therapy as first line treatment for mHSPC.** The CS estimates that most patients in England and Wales with mHSPC (around 70 – 80%) begin treatment with ARTA + ADT doublet therapy,

and around 10-20% of patients receive ARTA + ADT + chemotherapy triplet therapy. The current NICE recommended doublet therapies are apalutamide + ADT, and enzalutamide + ADT. The only triplet therapy currently recommended by NICE is darolutamide + docetaxel + ADT. The EAG's clinical expert commented that doublet therapy is beneficial for most patients and the minimum standard of care is to consider at least doublet therapy for patients who are fit enough. The expert explained that triplet therapy (a 'stronger' regimen due to the inclusion of chemotherapy) would be considered the preferred option in patients who have poor prognostic features at diagnosis who may not respond adequately to doublet therapy. Typically, these would be patients with high volume disease, or visceral metastasis. The expert commented that the decision to give triplet therapy can be influenced by the patient's pathology, for example if they have a Gleason primary pattern 5 cancer. This is the highest grade in the Gleason grading system for prostate cancer, indicating the most aggressive and poorly differentiated cancer cells. The company's experts advised that if a patient is fit enough to receive chemotherapy, they would be offered triplet therapy.

- The CS, in the EAG's interpretation, appears to suggest that triplet therapy is the preferred standard of care in mHSPC and would be given to all patients who can tolerate, and are willing to undergo, chemotherapy. The EAG's clinical expert had a slightly different view, commenting that in her experience triplet therapy tends to be targeted to patients with a disease pattern that demonstrates poorer prognosis features at diagnosis. Another consideration mentioned by the clinical expert is that if first line treatment includes docetaxel it is unlikely that docetaxel would be given as a subsequent treatment when the cancer progresses. Some patients prefer to begin their treatment with an ARTA + ADT regimen, and reserve docetaxel as a possible future treatment option when their ARTA treatment response attenuates (assuming they will still be able to tolerate chemotherapy later on).
- **The company proposes darolutamide + ADT as a first line treatment option in mHSPC, specifically for patients ineligible for docetaxel.** The CS notes that some patients are unable or unwilling to tolerate the cytotoxic effects of docetaxel chemotherapy. The docetaxel-ineligible patient population has been considered in previous NICE technology appraisals, namely TA741 (apalutamide plus ADT in mHSPC), TA412 (Radium-223 dichloride in mCRPC), and in an NHS England Clinical Commissioning Policy Statement (in mHSPC).[1] In essence, suitability for docetaxel is made on an individual patient basis informed by a risk-benefit assessment of various patient factors.

- We discuss the company's proposed restriction in their decision problem to docetaxel ineligible patients (see section 3). The CS suggests that there is current unmet need for a non-chemotherapy doublet treatment regimen in mHSPC. Although apalutamide + ADT and enzalutamide + ADT are established agents which can be used in patients who are deemed unsuitable for docetaxel, the company argues that they are associated with significant treatment-related toxicities (e.g. central nervous system adverse effects, fatigue, hypertension, seizures, skin toxicity) and drug-drug interactions. Darolutamide, in contrast, is claimed to have a more favourable tolerability profile and is less likely to interact with other medications including those used to treat common comorbidities (e.g. cardiovascular disease). The company provides a detailed justification in support of their view, citing their expert advisory board, clinical trial data and an APCCC consensus opinion. The EAG's clinical expert is of the opinion that there isn't necessarily unmet need in mHSPC but acknowledged that enzalutamide and apalutamide are associated with certain adverse effects, notably fatigue, central nervous system effects (e.g. seizures, risk of falls) and rash. In her experience apalutamide tends to cause more adverse effects and consequently it is given less often compared to enzalutamide. Darolutamide, based on her clinical experience of prescribing triplet therapy, is well tolerated.

EAG comment on the background information

The CS provides detailed and comprehensive background information on mHSPC, the current care pathway, and the proposed use of darolutamide + ADT as an initial treatment option for patients unsuitable for docetaxel in the metastatic disease setting. The EAG's clinical expert generally agreed with the company's description of the current care pathway though noted likely variation in clinical practice. The company suggests that darolutamide + ADT offers potential advantages over currently used ARTA + ADT treatments, including better tolerability and fewer adverse effects. The EAG's clinical expert advisor concurs with this, based on clinical experience of treating patients with darolutamide triplet therapy. These advantages are attributed to the distinct chemical structure of darolutamide. The EAG's view is that despite its novel features, darolutamide can be regarded as sufficiently similar in mechanism of action to the current NICE recommended ARTAs for mHSPC (i.e. apalutamide + ADT and enzalutamide + ADT). This is one of the factors necessary to support the case for a cost-comparison rather than a cost-utility technology appraisal.

3 CRITIQUE OF THE COMPANY'S DEFINITION OF THE DECISION PROBLEM

Table 3 below summarises the decision problem addressed by the company in the CS in relation to the final scope issued by NICE and the EAG's comments on this. Overall, the company's decision problem matches the NICE scope, with a couple of exceptions:

- **Population: docetaxel-ineligible.** The company specify a docetaxel-ineligible population in their decision problem, whereas the population in the NICE scope is broader (all people with mHSPC) which aligns with the indication in the draft Summary of Product Characteristics (SmPC)[2] for darolutamide. However, the comparators in the NICE scope are described as "for people in whom docetaxel is not suitable" which implies that a docetaxel-ineligible population is relevant to this appraisal. The EAG do not view this as a decision problem issue.
- **Subgroups: none**. The company chose not to assess darolutamide + ADT in the two subgroups in the NICE scope: patients with de novo (newly diagnosed) metastatic disease and patients with high-risk disease. We find that the rationale for the similarity of results of the de novo group compared to the ITT population would be better backed up with data, however the other justifications are appropriate (see Table 3 below).

Table 3 Summary of the decision problem

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
Population	People with hormone-sensitive metastatic prostate cancer	Adult men with mHSPC who are unsuitable for chemotherapy.	The NICE scope matches the proposed licensed indication for darolutamide + ADT which is treatment of adult men with mHSPC (draft SmPC 4.1). The company's decision problem restricts the population to adult men with mHSPC who are unsuitable for chemotherapy, i.e. docetaxel-ineligible. The company states that this population aligns with the chosen cost-comparator, apalutamide + ADT, which NICE recommends as an option for mHSPC in patients who are unsuitable for docetaxel (CS section B.1.1) and the EAG agrees that this is appropriate. As this population aligns with the definition of the comparators in the NICE scope, as treatments for people in whom docetaxel is not eligible, there is no issue.
Intervention	Darolutamide with androgen deprivation therapy	Darolutamide with androgen deprivation therapy.	As per scope; no comment.
Comparators	For people in whom docetaxel is not suitable:	Apalutamide with androgen deprivation therapy.	The choice of apalutamide + ADT as the comparator is in accordance with the NICE criteria for cost comparisons. Enzalutamide + ADT also meets the NICE criteria for cost

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
	• Apalutamide and androgen deprivation therapy • Enzalutamide and androgen deprivation therapy	As this submission is a cost comparison, we have compared darolutamide with a single NICE-recommended comparator, apalutamide. NICE technology appraisal 741 recommends apalutamide at the same point in the treatment pathway with the same wording as is anticipated for darolutamide. That is, apalutamide is recommended for people with mHSPC who are unsuitable for chemotherapy.	comparisons. The rationale provided by the company is that enzalutamide is not used in the same docetaxel ineligible population as proposed for darolutamide + ADT (clarification response A1), although the NICE TA712 recommendation states that enzalutamide + ADT offers another option for people who cannot have docetaxel.[3] In the EAG's view, either comparator would be appropriate for the population in the NICE scope and in the indicated population in the SmPC. The company is permitted to include just one cost comparator treatment, or more than one if preferred.
Outcomes	The outcome measures to be considered include: • Overall survival • Progression-free survival • Response rate • Time to hormone relapsed prostate cancer	The outcome measures to be considered include: • Overall survival • Radiographic progression free survival • Time to castration resistant prostate cancer • Time to subsequent therapy • Prostate-specific antigen undetectable rate • Time to prostate-specific antigen progression	The company decision problem includes all outcomes in the scope except response rate for which the company justification is appropriate.

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
	• Time to subsequent treatment • Prostate-specific antigen undetectable rate • Time to prostate-specific antigen progression • Time to pain progression • Adverse effects of treatment • Health-related quality of life.	• Time to pain progression • Adverse effects of treatment • Health-related quality of life. • **Radiographic progression free survival (rPFS)** was the primary endpoint in the ARANOTE study. • **Response rate** was not a pre-planned endpoint in the ARANOTE study and thus these data will not be included in this submission. Response rate is not generally used as an outcome measure in advanced prostate cancer, as prostate metastases, particularly bone metastases, generally do not show radiological responses to treatment, even though overall the treatment may be working.	
Economic analysis	The NICE reference case stipulations for expressing cost-effectiveness in ICERs, cost-comparisons, time horizon, cost	Cost-comparison model considered from an NHS perspective.	The company provided a simple cost-comparison analysis that evaluates the difference between the drug acquisition costs of darolutamide + ADT and apalutamide + ADT. The company assume that everything else, mortality, administration costs, resource use, adverse events, etc., are the same.

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
	considerations, commercial arrangements, and availability of biosimilars, etc., should be taken into account. [Abridged version of the text in CS Table 1.]		There is no function available in the cost-comparison spreadsheet to verify or explore differing parameters (clarification question B1). The cost-comparison uses a lifetime horizon, and the costs are considered from an NHS and PSS perspective, which are both appropriate. Further details are in section
Subgroups	If the evidence allows, the following subgroups of people will be considered: • people with newly diagnosed metastatic prostate cancer • people with high-risk metastatic prostate cancer	No subgroups. **Adult men with newly diagnosed metastatic prostate cancer** Both patients with M1 (de novo) and M0 (recurrent) at initial diagnosis have been included in ARANOTE. The majority of patients (72.5%) were de novo and the results in ARANOTE have been consistent across these subgroups. Therefore, the appraisal has focused on the ITT population. Consistency between these subgroups gives further re-assurance that darolutamide is similarly efficacious in both	No subgroups. De novo disease is reported in the ARANOTE pivotal trial as a baseline characteristic (metastases at initial diagnosis: de novo/recurrent/unknown; CS Table 8) and de novo participants comprise the majority (72.5%). It is not possible for the EAG to verify whether their results are consistent with the ITT population because they are not included in the results of the pre-specified subgroup analyses (CS Figure 10). Results from all the other pre-specified subgroup analyses are consistent with the results from the ITT analyses, however it would have

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
		newly diagnosed de novo patients and patients with mHSPC in general. **Adult men with high-risk metastatic prostate cancer** It is not clear what the high-risk metastatic prostate cancer definition is in the scope. ARANOTE has been stratified by extent of disease (i.e. non-regional lymph node metastasis, bone metastasis, and visceral metastasis). The efficacy observed in ARANOTE was consistent across these three subgroups. There was no classification by 'high-risk' disease in ARANOTE. There is inconsistent use of 'newly diagnosed' and 'high risk' for randomisation across all mHSPC trials. Furthermore, although appraisals for apalutamide in mHSPC also listed these subgroups in their scopes they were never explored by the submitting company nor was the lack of data in these subgroups highlighted as a key issue during the	been useful for the company to present the de novo group alongside these to verify their statement. We find that high-risk mHSPC is difficult to define. Our clinical expert advised that definitions of high-risk differ between metastatic and non-metastatic disease. For metastatic disease risk is based on extent of bone metastases or presence of visceral metastases. In non-metastatic disease high risk is based on the Cambridge prognostic score. There is no specific subgroup in the pivotal ARANOTE trial that represents 'high-risk' mHSPC. The pre-specified subgroups in the ARANOTE trial cover various markers of high-risk, e.g. high-volume disease, presence of visceral metastases, high Gleason score ([redacted]8), but are not definitions of high-risk disease in themselves. Results for these individual subgroups, however, were also generally consistent with the results of the ITT population.

	Final scope issued by NICE	Company's decision problem and rationale if different from the final NICE scope	EAG comments
		appraisal. As such, this appraisal has focused on the ITT population.	
Special considerations including issues related to equity or equality		[The company considers the following: • patients with a history of seizures • patients with multiple comorbidities at risk of drug-drug interactions (DDIs)] • [Rationale is in the full text of the company decision problem in CS Table 1.]	**Patients with a history of seizures.** Darolutamide is unique among ARTAs in that it does not cross the blood brain barrier (see also section 2.2.2). This is a small patient group (ARANOTE pivotal trial n=1; ARASENS trial n=6; EAG's clinical expert has seen 1 patient (on enzalutamide) with seizures in the last 5 years) that would benefit from the addition of darolutamide + ADT as a treatment option. **Patients with multiple comorbidities.** The number of DDIs associated with darolutamide are significantly fewer than for other ARTAs.[4-6] The availability of treatment with darolutamide may improve ease of medication management for mHSPC patients who are frequently already receiving multiple drugs.

Source: Reproduced from CS Table 1 with some abridgement, and additional EAG comments.
Abbreviations: ADT, androgen deprivation therapy; ARTAs, androgen receptor targeted agents; DDIs, drug-drug interactions; ICERs, incremental cost-effectiveness ratios; ITT, intention to treat; mHSPC, metastatic hormone sensitive prostate cancer; rPFS, radiographic progression free survival.

EAG comment on the company's decision problem

The company's decision problem is similar to the NICE scope for this technology appraisal, with the main difference being company's decision not to assess cost comparison for the de novo (newly diagnosed) and high-risk patient subgroups in the NICE scope. The EAG notes that the other current NICE recommended treatment for patients with mHSPC, enzalutamide + ADT, is also used to treat patients unsuitable for docetaxel. The company favoured apalutamide + ADT as their chosen cost comparator treatment, but did not explicitly state whether enzalutamide + ADT would also be an appropriate comparator. The company is permitted to select just one NICE recommended treatment for comparison with darolutamide + ADT, but also has the option of comparing against more than one recommended treatment.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company conducted a systematic literature review (SLR) to identify randomised controlled trials and real-world evidence of the clinical effectiveness of treatments for mHSPC (CS Appendix D). An SLR for studies reporting cost and healthcare resource use for mHSPC was also conducted (CS Appendix G); but is not referred to in the main submission document.

The clinical effectiveness SLR methods are mostly adequate (a summary of the EAG's appraisal is in Appendix 1). The clinical effectiveness SLR included 42 studies overall (CS Appendix Table 15). Relevant to this appraisal, the company identified one RCT that evaluated darolutamide + ADT compared to placebo + ADT in mHSPC: the pivotal phase III ARANOTE trial[7] which is discussed in section 4.2 below, and one RCT that evaluated apalutamide + ADT compared to placebo + ADT in mHSPC: the phase III TITAN trial,[8] discussed in section 4.3 below.

EAG comment on the review methods

Minor aspects of reporting the methods are missing from the CS, and the company may not have provided the correct excluded studies list, but on investigation we found that no relevant studies with results have been omitted. We agree that ARANOTE and TITAN are the included studies that provide the most relevant results for this appraisal.

4.2 Critique of the ARANOTE trial

The ARANOTE trial is a company-sponsored international phase III randomised placebo-controlled trial evaluating the clinical efficacy and safety of darolutamide + ADT vs placebo + ADT.[7]

4.2.1 ARANOTE study design

Table 4 below summarises the ARANOTE trial study design.

Table 4 Overview of the ARANOTE trial

Study characteristic	Details
Study design	RCT. Randomised intervention: placebo 2:1; stratified according to presence of visceral metastases and use of prior local therapy. Double blind until primary completion analysis; open label thereafter.

Study characteristic	Details
Location	133 sites in 15 countries across Europe, Asia, Australia and the Americas. No UK sites or patients.
Population	Men with mHSPC, including both de novo disease, i.e. metastatic at diagnosis, (72.5%) and recurrent disease (approximately 21%).
Pre-specified subgroups	Age (<65/65-74/75-84/≥85), PSA (<median/≥median), ECOG PS (0/≥1), Gleason score (<8/≥8), disease volume (high/low), race (White/Asian/Black/other), Region (Europe and rest of the world/Asia/Latin America), Visceral metastases (yes/no), prior local therapy (yes/no).
Key eligibility criteria	• Confirmed metastatic adenocarcinoma of the prostate • ECOG PS of 0, 1 or 2 • Started ADT ≤12 weeks before randomisation • Adequate bone marrow, liver and renal function • Prior chemotherapy (docetaxel or immunotherapy) for prostate cancer was not permitted.
Intervention	Darolutamide (600 mg BID) + ADT (n=446)
Comparator	Placebo (darolutamide matched tablets BID) + ADT (n=223) NB in practice this represents ADT monotherapy which is no longer standard of care in the NHS (see section 2.2.3).
Primary outcome	**Radiological progression free survival (rPFS**) (see section 4.2.4 and 4.2.5.1)
Secondary outcomes	**OS** (see section 4.2.4 to 4.2.5.2) Time to initiation of subsequent cancer therapy Time to CRPC Time to PSA progression PSA undetectable rate Time to pain progression **Adverse events** (see section 4.2.4 and 4.2.6)
Other outcomes	PFS2 (investigator-assessed) Time to symptomatic skeletal event **Time to deterioration in FACT-P total score** (see section 4.2.4) Time to first prostate cancer-related invasive procedure
Crossover	• [redacted] participants who were still on study treatment in the placebo + ADT arm crossed over to open label darolutamide + ADT after the primary completion analysis (board approved, due to ethical reasons). • RPSFT and IPE statistical methods (pre-specified in the SAP) were used to adjust for treatment switching in sensitivity analyses of OS (see section 4.2.4).
Duration of study	After a 28-day screening period, participants commenced treatment with the study drug and were assessed in clinic every 12 weeks for 12

Study characteristic	Details
	± 1 months. Thereafter, participants were contacted every 12 weeks until death, loss to follow-up, withdrawal of consent or end-of-study.
Main analyses	**Primary completion analysis:** assessed primary outcome of rPFS and all other outcomes; median follow-up: 25.3 months and 25.0 months for darolutamide + ADT and placebo + ADT respectively; database lock June 2024. **Final OS analysis:** assessed OS and safety; median follow-up: [redacted] according to original assignment to darolutamide + ADT and placebo + ADT arms respectively; database cut-off [redacted]

Source: CS section B.3.3.1; CS Figure 10; CS Tables 7 and 8; Saad 2024[7]; Final OS Results Summary[9].
Abbreviations: ADT, androgen deprivation therapy; BID, bis in die [twice a day]; CRPC, castration-resistant prostate cancer; EAG, evidence assessment group; ECOG PS, Eastern Cooperative Oncology Group Performance Status; IPE, iterative parameter estimate; mHSPC, metastatic hormone-sensitive prostate cancer; OS, overall survival; PFS, progression free survival; PSA, prostate-specific antigen; RCT, randomised controlled trial; rPFS, radiographic progression free survival; RPSFT, rank preserving structural failure time; SAP, statistical analysis plan; UK, United Kingdom.
Outcomes in **bold font** are used in the company's indirect treatment comparison (CS section B.3.9.3; section 4.3 of this report).

EAG comment on the ARANOTE trial design

The ARANOTE study is a generally well-designed RCT. The EAG doesn't have any major concerns about the study or its appropriateness to inform this NICE technology appraisal.

4.2.2 ARANOTE population baseline characteristics

Participant demographic and clinical characteristics at baseline are reported in CS Table 8. We agree with the company assessment of the baseline characteristics of participants in the ARANOTE trial. All characteristics are well balanced between arms, and therefore any reported prognostic factors are well balanced too.

The company note the presence of more advanced disease than the general mHSPC population in this trial due to large proportions of participants with high Gleason scores, de novo disease, and high-volume disease (CS section B.3.3.2). The EAG's clinical expert also advised that the proportion of participants with visceral metastases was high at 12% (CS Table 8) compared to less than 5% in her clinical practice. However, the proportions of patients with these characteristics were balanced between arms and would not bias the trial results.

There are no UK participants in the ARANOTE trial, but our clinical expert confirmed that the baseline characteristics (apart from the high presence of visceral metastases) for the trial population are generally representative of the overall NHS mHSPC population in England and noted that the proportions of Asian and Black participants were representative for a trial. Our expert also noted the low median prostate specific antigen (PSA), approximately 21 ng/mL, whereas she might expect to see approximately 40 ng/mL in practice, however the overall range for serum PSA is wide which is representative.

The ARANOTE trial had broad eligibility criteria (CS section B.3.3.1) and did not prospectively recruit docetaxel-ineligible participants, although prior chemotherapy for prostate cancer was an exclusion criterion. Our clinical expert advised us that there is no strict definition of docetaxel ineligibility and that clinicians make a risk-benefit consideration for chemotherapy for each patient in practice, considering performance status, presence of peripheral neuropathy, diabetes and severity of cardiovascular disease. The baseline characteristics show that most of the participants had an Eastern Cooperative Oncology Group (ECOG) performance score of 0 or 1 compared to only 3-4% with an ECOG score of 2, and none with a score of 3 or 4 (according to trial eligibility criteria), which alongside inclusion criteria of adequate bone marrow, liver and renal function (CS Table 7) suggests a reasonably fit population.

In addition, the company note in clarification response A9 that the subgroup analysis results for age and ECOG performance status are consistent with the overall trial results. They state that over 91% of participants had at least one comorbidity upon study entry and that the most common comorbidities were vascular disorders, musculoskeletal and connective tissue disorders, renal and urinary disorders and metabolism and nutrition disorders. There is no data from ARANOTE on whether the participants would have chosen not to receive docetaxel if offered (which is the other apalutamide Blueteq criterion for docetaxel-ineligibility). However, the EAG considers that patients who choose not to receive docetaxel for any reason are likely to be representative of the general mHSPC population. On balance, it is likely that the population in the ARANOTE trial is appropriate to represent a docetaxel-ineligible mHSPC population but equally also includes patients suitable for docetaxel.

EAG comment on participant baseline characteristics

4.2.3 ARANOTE risk of bias assessment

The company used the “NICE checklist for RCTs” (which the EAG recognises as the criteria for appraising RCTs devised by the Centre of Reviews and Dissemination (CRD) for

systematic reviews), to judge the methodological quality of the trial. In their judgment the ARANOTE trial is at low risk of bias (CS section B.3.5 and CS Table 11).

The EAG appraised the study using the same checklist (see Appendix 2 of this report). For most of the questions our response agrees with that of the company – that the study is at low risk of bias. However, we introduced a distinction between risk of bias in the randomised double-blind phase and in the open label follow up study period. This distinction was applied to two questions where the risk of bias potentially changes over the course of a study.

For the question 'Were the care providers, participants and outcome assessors blind to treatment allocation? 'we answered 'yes' (low risk of bias) for the primary outcome rPFS. This outcome was assessed only at the primary completion analysis of the double-blinded phase and was prior to study unblinding and the option to crossover from placebo to darolutamide. Blinded independent central review (BICR) used to assess rPFS according to standardised measures RECIST v1.1 and PCWG3 criteria of rPFS is a further justification for our judgment.

Outcomes reported after the primary completion analysis, that is, final OS and final adverse events we regard as at high risk of bias. This is due to the effects of unblinding (performance bias) and from patients subsequently crossing over from placebo + ADT arm to the darolutamide + ADT arm. We do, however, acknowledge the company's view that the impact of crossover on OS can be considered reduced given that:

- the period of crossover for the final OS analysis was [redacted], and
- of the [redacted] randomised placebo patients still on study treatment who crossed over to darolutamide after primary completion, only [redacted] died under the darolutamide crossover period (to put this into context, there were [redacted] deaths reported in the final OS analysis after primary completion ([redacted] in the darolutamide arm, [redacted] in the placebo arm)).

EAG comment on risk of bias

The pine characteristics are well balanced between trial arms, and are generally representative of the overall mHSPC population in the NHS. The ARANOTE trial did not specifically recruit docetaxel-ineligible participants, however the trial results are likely generalisable across the mHSPC patient spectrum.

Primary outcome of rPFS, and outcomes assessed at the primary completion analysis, are at low risk of bias, however OS and adverse event outcomes that

were assessed after the primary completion analysis (clarification response A3) are at high risk of performance bias and unclear risk of attrition bias.

4.2.4 Outcomes assessment

The ARANOTE trial outcomes incorporated into the cost comparison model are time on treatment and rPFS (from the darolutamide + ADT arm of the trial only) (see section 4.3 of this report). Overall survival, FACT-P, and adverse events are not included in the model.

Radiological PFS (rPFS) was the primary outcome of the ARANOTE trial, defined as time from randomisation to radiological progressive disease in soft tissue (RECIST v1.1 criteria) or bone (PCWG3 criteria), or all-cause death (CS Table 7). It was assessed every 12 weeks by BICR (CS Table 7), and results were reported for the primary completion analysis which had a median follow up of approximately 25 months which, our clinical expert confirms, is adequate time to assess response (CS section B.3.6.2).

rPFS is used as a surrogate outcome for OS in a sensitivity analysis of the company's ITC (CS section B.3.9.4.1 and section 4.3.6.2 of this report).

Overall survival (OS) was a key secondary outcome, defined as time from randomisation to all-cause death (CS Table 7). It was the first secondary outcome in the hierarchy for testing of statistical significance (CS Table 10) and as the OS results were not statistically significant (section 4.2.5.2 below) no further outcomes were tested for statistical significance in the ARANOTE trial. At the final analysis the median OS was [redacted] in either treatment arm (CS section B.3.6.3.1).

The results for the final overall survival analysis may be subject to confounding due to crossover of [redacted] participants from the placebo + ADT arm to the darolutamide + ADT arm. The EAG believe that it is appropriate to carry out adjustment for crossover because the crossover in the trial does not reflect the treatment pathway in clinical practice, although the company view the impact of crossover as minimal because there were only [redacted] deaths among the darolutamide crossover participants in the brief [redacted] crossover period (clarification response A5).

Published statistical methods, pre-specified in the trial's statistical analysis plan, were used to adjust for patient crossover in the ARANOTE trial following unblinding at the primary completion analysis. Two approaches were considered, the rank preserving structural failure time (RPSFT) and iterative parameter estimate (IPE) methods. These methods estimate the treatment effect as if patients in the placebo arm had never crossed over to darolutamide.

The company justify their selection of these methods in clarification response A6. They note that the RPSFT and IPE methods are two well-known methods used to adjust for crossover in randomised trials and are commonly used in NICE appraisals. They acknowledge that other methods can be used but note that "all such methods are subject to limitations". They do not elaborate on such limitations, including any applicable to their chosen methods (i.e. RPSFT and IPE). A discussion of the merits and limitations of the available adjustment methods in relation to the ARANOTE trial, including consideration of clinical plausibility would have given a stronger rationale for the company's selected adjustment methods.

The EAG invited the company to provide cross-over adjusted OS results using other available methods (e.g. featured in NCE DSU TSD number 16) if available (clarification question A7). The company responded that only the RPSFT and IPE methods were performed as there were very few OS events during the crossover period (between primary completion and final OS analysis). The company point out that the similar OS estimates the from RPSFT and IPE methods "provides confidence that the appropriate methodologies have been applied for these analyses." (clarification question A6). Whilst consistency in results is reassuring, it is only one consideration in choosing an appropriate analysis method. The EAG would have preferred to see of crossover adjusted OS estimates from all available methods to assess the degree to which they are consistent, as this would provide a more informed consideration of which adjustment methods, if any, are appropriate to inform decision making.

Time to deterioration in the Functional Assessment of Cancer Therapy – Prostate (FACT-P) total score is a pre-specified outcome (not primary or secondary) in ARANOTE which is used in the company's ITC. FACT-P is a validated patient reported outcome measure for patients with prostate cancer.[10] The ARANOTE trial defined deterioration as a decline of $\geq$10 points from baseline in the total score. This is a conservative use of the published estimate of clinically meaningful change that is six to 10 points in total score change;[11] and the time to deterioration is measured from randomisation (Clinical study report (CSR) section 5.1.4.3).

Adverse events are reported for the Safety Analysis Set (SAF) which consisted of all participants who received $\geq$1 dose of the study drug and participants were analysed according to the study drug they received (CS Table 9). Namely: darolutamide + ADT (double blind period), darolutamide + ADT (double-blind and open-label periods), placebo + ADT (double-blind period), placebo-darolutamide (crossover, i.e. open-label, period). Adverse events are reported according to the NCI-CTCAE v 5.0 criteria, and the results are

reported from the Final Analysis (CS section B.3.10). A pre-specified analysis of exposure adjusted incident rates (EAIRs) is also reported for ARTA-related adverse events, such as hypertension, flushing, diabetes, fatigue, and rash, during the double-blind study period (full list in CS Table 30; also considered as adverse events of special interest in CS Table 29).

EAG comment on outcomes assessment

The outcome measures included in the ARANOTE trial are similar to outcomes commonly used in oncology clinical trials, including rPFS and OS which were considered in the NICE technology appraisal of apalutamide + ADT in mHSPC (TA741). Caution is advised when interpreting the final OS estimates because the data are immature and subject to confounding due to crossover. Crossover adjusted OS estimates are available but do not necessarily represent estimates from alternative crossover adjustment methods.

4.2.5 Key efficacy results of the ARANOTE trial

All results (except for safety) are reported for the full analysis set (FAS), that is all randomised participants according to the treatment arm they were allocated at randomisation (CS Table 9) equivalent to an intention-to-treat (ITT) analysis. Safety results, including adverse events, are reported for the safety analysis set (SAF), see section 4.2.6.

4.2.5.1 Radiological Progression Free Survival (rPFS) – primary outcome

The results for rPFS are statistically significant and in favour of treatment with darolutamide + ADT (CS section B.3.6.2):

- At the primary completion analysis (after 222 events; data cut off 7 June 2024), participants treated with darolutamide + ADT had a 46% reduced risk of rPFS or death compared to participants in the placebo + ADT arm (HR, 0.54; 95% CI, 0.41 to 0.71; p<0.0001).
- At 24 months (within the double-blind period), the rPFS rate was 70% in the darolutamide + ADT arm compared to 52.1% in the placebo + ADT arm (median rPFS was not reached in the darolutamide + ADT arm).
-

Results for all the pre-specified subgroups were consistent with the results for the full analysis set (FAS). The subgroups for age ≥85, Black race, and presence of visceral metastases have wide confidence intervals due to small sample sizes (CS section B.3.7; CS Figure 10). The results for subgroups that could indicate high-risk disease, e.g. Gleason

score ≥8, high volume disease, presence of visceral metastases, are supportive of the positive effect of darolutamide + ADT.

4.2.5.2 Overall survival (OS) – key secondary outcome

Overall survival results are summarised in Table 5 below.

Table 5 ARANOTE overall survival results (darolutamide + ADT vs placebo + ADT; FAS)

Analysis	Risk reduction	Hazard Ratio	95% CI	p-value
Primary completion analysis (ITT) (163 events; data cut-off 7 June 2024)	19% [a]	0.81	0.59 to 1.12	0.1007
Final analyses				
ITT ([redacted] events; data cut-off [redacted])	[redacted]	[redacted]	[redacted]	[redacted]
RPSFT sensitivity analysis (to adjust for crossover)	[redacted] [a]	[redacted]	[redacted]	[redacted]
IPE sensitivity analysis (to adjust for crossover)	[redacted] [a]	[redacted]	[redacted]	[redacted]

Source*:* CS section B.3.6.3.1
Abbreviations: CI, confidence interval; FAS, full analysis set; IPE, iterative parameter estimate; ITT, intention to treat; RPSFT, rank preserving structural failure time.
[a] calculated by EAG.

Results for OS were not statistically significant, but they show a positive trend in favour of treatment with darolutamide + ADT; median OS was [redacted] in either treatment arm; sensitivity analyses to adjust for crossover to the darolutamide + ADT arm were consistent with the results from the final analysis (CS section B.3.6.3.1). Results for all pre-specified subgroups are [redacted] of darolutamide + ADT treatment (hazard ratios range from [redacted]), however as the confidence intervals for all subgroups [redacted] in the forest plot reported in Figure 5-2 of the Final Overall Survival Results Summary[9] the subgroup results are highly uncertain.

4.2.5.3 Other outcomes

Results for other secondary outcomes are also favourable to treatment with darolutamide + ADT compared to placebo + ADT, and they are reported in CS sections B.3.6.3.2 to B.3.6.3.6.

Results for time to deterioration in FACT-P, an outcome used in the comparative effectiveness NMA, are not reported in the CS. At the primary completion analysis, participants treated with darolutamide + ADT had an approximately [redacted]% reduced risk of deterioration in FACT-P total score compared to participants treated with placebo + ADT (HR [redacted]; 95% CI [redacted]; p[redacted]) (CSR section 5.1.4.3).

4.2.6 Key safety results of the ARANOTE trial

Safety results are reported in CS section B.3.10, with topline results reported in the Final Overall Survival Results Summary.[9]

A summary of treatment-emergent adverse events (TEAEs) in ARANOTE is reported in CS Table 27. It shows the rates for TEAEs are very similar across the darolutamide and placebo treatment arms for participants experiencing any adverse events, serious adverse events, Grade 3 or 4 adverse events, or Grade 5 adverse events, and AEs leading to permanent discontinuation. Slightly fewer dose modifications were reported in the placebo arm. The lower rates in the placebo-to-darolutamide group are explained by the shorter time frame of the post-crossover period.

The most common any-Grade TEAEs experienced by 10% or more participants were anaemia, arthralgia, urinary tract infection and back pain: the proportion of participants experiencing these in the darolutamide arm was almost identical pre- and post-crossover, and the proportion of participants in the placebo arm (double-blind period only) was slightly lower and, for back pain, about the same (CS section B.3.10.1).

The most common Grade 3 and 4 TEAEs in 5% or more participants were hypertension, anaemia, increased aspartate aminotransferase (AST), increased alanine aminotransferase (ALT) and bone pain (CS section B.3.10.1). These are all adverse events related to ARTA treatments, but numbers of events were similar between the darolutamide and placebo arms.

TEAEs of special interest are those related to ARTA treatments. CS Table 29 shows that proportions of participants experiencing these adverse events was low ([redacted]) and they are similar between darolutamide and placebo arms. For fatigue and asthenia (weakness) rates are lower in the darolutamide arm compared to the placebo arm. When the results are adjusted for exposure (double blind trial period only), the incidence risk ratios for experiencing adverse events are lower in the darolutamide arm compared to the placebo arm not only for fatigue, but also for hypertension, diabetes mellitus, decreased weight, heart failure, depressed-mood disorder, and cerebral ischemia (CS Table 30).

EAG comment on safety results

Safety results from ARANOTE show darolutamide + ADT has a similar safety profile to placebo + ADT. For fatigue and weakness darolutamide + ADT was shown to be better than placebo + ADT. Fatigue and weakness do not tend to require hospitalisation and therefore these improvements do not necessarily incur cost savings, however the EAG's clinical expert advised that improvement relating to fatigue and weakness are of immense value to the patient.

4.2.7 Pairwise meta-analysis of intervention studies

No pairwise meta-analysis was conducted as there is only one relevant included study with results, the ARANOTE trial. The EAG concurs with the CS that a pairwise meta-analysis is currently not possible.

4.3 Critique of the indirect treatment comparison (ITC)

4.3.1 Rationale for ITC

The company's rationale for conducting an indirect treatment comparison (ITC) is based on the lack of direct, head-to-head evidence comparing darolutamide + ADT with the company's chosen cost-comparison treatment, apalutamide + ADT. The EAG agrees that an ITC is necessary to address the decision problem.

4.3.2 Identification, selection and feasibility assessment of studies for ITC

The CS reports that an ITC comparing darolutamide versus apalutamide (both in combination with ADT) is possible because both treatments have been compared to placebo + ADT in clinical trials (CS section B.3.9). There were two relevant placebo-controlled trials available for inclusion in the ITC, the TITAN trial[8 12] (apalutamide + ADT versus placebo + ADT) and the aforementioned ARANOTE trial[7] (darolutamide + ADT versus placebo + ADT). TITAN was the pivotal phase III multi-centre RCT which supported the regulatory approval of apalutamide + ADT in mHSPC and which informed NICE's recommendation for apalutamide as an option for treating mHSPC (TA741)[13] in 2021.

Although not explicitly labelled as such, the CS reports a feasibility assessment of the TITAN and ARANOTE trials as evidence to inform the ITC, considering factors such as comparability of the trial designs, the availability of outcome measure data, compatibility of outcome definitions, and methodological quality and risk of bias. (CS section B.3.9 and CS Appendix D Tables 16-20; clarification response A10).

4.3.3 Clinical heterogeneity assessment

In terms of patient characteristics, the CS compared the trials on factors including age, ethnicity, ECOG performance status, Gleason score, de novo disease, high volume disease and visceral metastases (CS Table 14). The CS comments that the trials are similar in terms of overall baseline characteristics, with exceptions for White ethnicity (around 10-13~~%~~ percentage points higher in TITAN), ECOG status 0 (higher in TITAN) and visceral metastases (just under 10 percentage points higher in ARANOTE). The EAG notes further differences between the trials not commented on in the CS, specifically high-volume disease (just under 10 percentage points higher in ARANOTE), and de novo disease (around 7 percentage points higher in TITAN). Overall. the above differences between the trials suggests that more patients in the ARANOTE trial have characteristics associated with worse prognosis and a potentially inadequate response to treatment than is the case for the TITAN trial population. This may potentially confound the results of the ITC, in favour of apalutamide + ADT. However, the magnitude of the differences in patient characteristics between TITAN and ARANOTE is relatively small (i.e. up to 10 percentage points difference) and thus any bias arising is unlikely to be substantial.

Most of the patient characteristics assessed in CS Table 14 are known prognostic factors and/or treatment effect modifiers in prostate cancer, though the CS does not explicitly identify them as such. The EAG invited the company to expand the list of patient baseline characteristics considered in CS Table 14, specifically to include any additional prognostic factors and effect modifiers (clarification question A11). The company responded that the list covers all key aspects and there were no further characteristics to add. Expert clinical advice to the EAG is that additional patient characteristics which should be considered include bone metastases (>4), PSA levels, presence of liver or lung metastases and haemoglobin and neutrophil counts.

4.3.4 Risk of bias assessment for studies included in the ITC

The company's assessment of bias for the ARANOTE trial is reported in CS Section B.3.5 (CS Table 11). As we have commented earlier (see section 4.2.3) we agree with the company that the trial is low risk of bias for outcomes measured at the primary analysis (prior to unblinding and patient crossover) but outcomes measured after primary analysis, i.e. OS and adverse events will be at increased risk of bias from performance bias and the effects of crossover from placebo + ADT to darolutamide + ADT. This is addressed in the CS using statistical adjustment methods to adjust for crossover (i.e. RPSFT and IPE methods), as discussed earlier in section 4.2.4.

The CS provides a risk of bias assessment for the TITAN trial using the University of York CRD critical appraisal criteria for RCTs (CS Appendix D1.3, Table 20). The company's conclusion is that "low risk of bias was found". The EAG has independently critically appraised the TITAN study using the same criteria and agrees with the company's judgement of low risk of bias overall. We previously assessed TITAN as at low risk of bias in in the EAG report for NICE TA741.[14] At that time interim results were available based on the double-blind randomised phase of the trial. Subsequently, the trial was unblinded and patients were permitted to crossover from placebo + ADT to apalutamide + ADT. A total of 208 of the 527 (39.4%) patients in the placebo + ADT arm crossed over to apalutamide + ADT. The final results of the TITAN trial are therefore potentially subject to high risk of performance bias and confounding because of crossover.[12]. However, as we note in the next section, a crossover adjusted estimate of OS from the trial has been reported.

4.3.5 Data inputs to the ITC

Separate evidence networks were constructed to estimate the relative effectiveness of the treatments for the following outcomes: rPFS, OS, time to deterioration in FACT-P, Grade 3-5 adverse events and discontinuation due to adverse events.

4.3.5.1 ARANOTE data inputs

The OS and adverse events analyses use ARANOTE data from the final OS analysis data cut-off from [redacted]. The rPFS and FACT-P analyses use data from the ARANOTE primary completion analysis data cut-off June 7th 2024. The rPFS and FACT-P outcomes were not intended to be updated at the [redacted] data cut, hence they have a slightly shorter median follow-up than the OS and adverse events analyses (median follow-up around 25 months compared to around [redacted], respectively).

The OS data used to inform the ITC were based on the ARANOTE full analysis population (analogous to a true ITT analysis) and not the crossover adjusted OS data reported in the CS (see section 4.2.5.2 of this report). Hence, the OS estimates for the placebo + ADT arm will be potentially confounded by survival outcomes in placebo patients who switched to darolutamide + ADT following study unblinding at the primary analysis in the trial. We discuss the implications of this for the results of the ITC below.

4.3.5.2 TITAN data inputs

The company clarified the source of the data inputs from the TITAN trial in clarification response A13. For the rPFS and FACT-P analyses, data from the primary analysis of TITAN [8] were used in the ITC, and for OS, Grade 3-5 adverse events and discontinuation due to adverse events, follow-up data from the final analysis set were used. [12] Hence, rPFS and

FACT-P analyses are based on shorter median follow-up than the OS and adverse events analyses (median follow-up around 24 months compared to around 44 months, respectively). The EAG notes the difference in median follow-up between the primary and final analyses in TITAN is larger than that of the ARANOTE trial, 20 months versus around [redacted], respectively.

As we mentioned earlier (section 4.3.4), the final OS estimates from TITAN are subject to the effect of patient crossover from placebo + ADT to apalutamide + ADT when the study was unblinded following the interim analysis. The journal publication of final analysis results from the trial.[12] reported a pre-planned sensitivity analysis using the inverse probability of censoring weights (IPCW) method to adjust for crossover. Under the IPCW method patients who crossed over from placebo to apalutamide were censored at the time of crossover, while patients remaining in the placebo group were weighted to compensate for missing data. The bias introduced by this informative crossover was corrected by weighting each patient by the inverse of their predicted probability of not being censored at a given time. The probability of crossover was determined by each patient's baseline characteristics in a logistic regression model. OS was then analysed with the censored data set and observations weighted by the inverse of the predicted probability of censoring.[12].

The OS HR for apalutamide + ADT vs placebo + ADT decreased from 0.65 (95% CI 0.53 to 0.79) to 0.52 (95% CI 0.42 to 0.64) when the IPCW adjustment was applied. Accordingly, the reduction in the risk of death with apalutamide increased from 35% to 48%, respectively, when the crossover adjustment was used. The EAG notes that the IPCW method is the only method of crossover adjustment mentioned in the trial publication, with no explicit rationale given for its use compared to other available methods. It is therefore unclear how consistent the crossover adjusted OS HR 0.52 (95% CI 0.42 to 0.64) is to OS estimates based on other adjustment methods.

4.3.5.3 Patient crossover adjustments

Importantly, the EAG notes that the ITC does not use the crossover adjusted estimates from the ARANOTE trial or the TITAN trial. Instead, the unadjusted ITT HRs from both trials are used as input parameters in the NMA.

In both trials, OS will be potentially confounded by placebo patients who switched to the experimental treatment (i.e. darolutamide + ADT, or apalutamide + ADT) following study unblinding. In each trial, the relative survival effects of the experimental treatment versus placebo will be potentially underestimated as a consequence. The CS does not mention whether crossover-adjusted OS HRs were considered for inclusion in the ITC and does not comment on the implications of using unadjusted OS HRs on the results of the ITC. The EAG suggests that all other things being equal, the use of unadjusted OS HRs will not bias the indirect comparison of darolutamide + ADT vs apalutamide vs ADT since in both trials the direction of bias is expected to be the same (i.e. underestimating the effect of the experimental treatment). However, differences in the magnitude of the bias between the trials may have an impact on the ITC. Specifically, the proportion of placebo patients who crossed over in TITAN (n=208/527; 39.4%) was [redacted] than in ARANOTE [redacted] and median follow-up was [redacted] (44 months versus [redacted] respectively). The implication is that the effect of crossover is likely to be greater in TITAN than in ARANOTE, due to more placebo patients crossing over and a longer follow-up period.

Differences in the impact of crossover are evident by comparing crossover adjusted and unadjusted OS HRs from the two trials (

Table **6**). There is little difference between the crossover adjusted and unadjusted HRs in the ARANOTE trial. The CS attributes this to very few deaths during the short time period between completion and final OS analysis [redacted] The difference in adjusted and unadjusted OS in TITAN is more pronounced and shows a greater reduction in the risk of death with apalutamide + ADT when crossover is adjusted for (unadjusted OS HR 0.65 (0.53 to 0.79); IPCW adjusted OS HR 0.52 (0.42 to 0.64), albeit using a different adjustment method to the ARANOTE trial. Given the fact that only selected crossover adjusted methods have been presented instead of a broader range of methods, the EAG considers it prudent to use the unadjusted ITT estimates from both trials in the base case ITC. However, the relative effectiveness of apalutamide + ADT versus placebo + ADT is underestimated by using the unadjusted estimate from TITAN (i.e. HR 0.65 vs HR 0.52). In turn this likely underestimates the true difference between darolutamide + ADT and apalutamide + ADT when compared indirectly, making them appear more similar than they actually are. The EAG would have liked to have seen sensitivity analyses using the crossover-adjusted OS estimate from TITAN (the IPCW methods plus any other available adjustment methods) to ascertain whether any significant differences in favour of apalutamide + ADT are detected. If so, it would weaken the case for a cost comparison appraisal because darolutamide would be inferior to apalutamide, at least in terms of OS.

Table 6 Crossover adjusted and unadjusted final OS estimates from the ARANOTE and TITAN trials

Trial ID, Treatments compared	**Proportion placebo patients crossed over, n/N (%)**	**Median follow-up (months)**	**Crossover unadjusted ITT HRs (95% CI)[a]**	**Crossover adjusted HRs (95% CI)**		
				Cross over adjustment method		
				RPSFT	**IPE**	**IPCW**
ARANOTE Darolutamide + ADT vs placebo vs ADT	[redacted]	31.4	[redacted]	[redacted]	[redacted]	NR
TITAN Apalutamide +ADT vs	208/527 (39.4)	44.0	0.65 (0.53 to 0.79)	NR	NR	0.52 (0.42 to 0.64)

Trial ID, Treatments compared	Proportion placebo patients crossed over, n/N (%)	Median follow-up (months)	Crossover unadjusted ITT HRs (95% CI)[a]	Crossover adjusted HRs (95% CI)		
placebo vs ADT						

Source: Table created by the EAG based on information in the CS, clarification question responses and the TITAN final survival analysis publication.[12]
Abbreviations: ADT, androgen deprivation therapy; CI, confidence interval; HR, hazard ratio; IPCW, inverse probability of censoring weights; IPE, iterative parameter estimate; ITT, intention to treat; NR, not reported; RPSFT, rank preserving structural failure time.
[a] used in the company's ITC

4.3.6 Statistical methods for the ITC

The CS describes the ITC as a network meta-analysis (NMA) using a Bayesian generalised linear model framework, citing NICE Decision Support Unit (DSU) technical support document 2[15] (CS section B.3.9.1). The EAG notes this is a standard approach commonly used to conduct NMAs informing NICE technology appraisals. The NMA uses the relative treatment effects between each of the treatment comparisons in the network and estimates the surface under the cumulative ranking (SUCRA) and mean ranks of each treatment (apalutamide + ADT, darolutamide + ADT and placebo + ADT).

4.3.6.1 Random effects versus fixed-effects modelling

The CS states that both random effects and fixed-effect models were considered, and that the fixed-effect approach was the *a priori* preferred approach. The justification for this was the assumption that the random effects models would not converge because of lack of data, given that only two studies were included. Later in the CS it is reported the random effects models did converge (CS section B.3.9.3.1). The CS notes the lack of heterogeneity between the trials in patients' baseline characteristics, trial design, and outcome definitions as another reason for favouring a fixed-effect approach. The EAG considers this a reasonable justification, notwithstanding the minor differences in patient prognostic factors between the trials, as we discussed earlier (section 4.3.3).

Random effects and fixed-effect model fitting statistics using the deviance information criterion (DIC) are reported for each outcome measure (CS Tables 15, 17, 19, 21 and 23), and in each case the CS reports there were limited differences in the DIC between the random and fixed-effect models indicating that both fit the data well (CS Appendix D.1.4). Given the company's *a priori* preference they report fixed-effect model results in their base

case for all outcomes (CS section B.3.9.3). For transparency they also provide the NMA results based on random effects (CS Appendix D.1.7) and comment that these results are "aligned in conclusion". The EAG considers the CS has adequately reported and justified the approach to fixed versus random effects modelling and has no particular concerns.

4.3.6.2 Surrogacy analysis

The CS notes that a limitation of the NMA is that the ARANOTE trial was not statistically powered to detect a difference in OS between darolutamide + ADT and ADT + placebo (CS section B.3.9.2). The company therefore conducted a surrogate sensitivity analysis to predict OS, informed by guidance on evaluating surrogate endpoints from NICE DSU TSD number 20.[16] CS Appendix J gives a detailed account of the rationale for this analysis, and methods used to validate the surrogate outcome (rPFS). The key aspects of the process include conducting a systematic review of trials of all treatments for mHSPC, to examine the relationship between rPFS and OS; conducing a correlation meta-analysis; and use of simulation modelling via Markov Chain Monte Carlo simulation.

The EAG considers that the company have provided good transparency in reporting their implementation of the method. The analysis confirmed that rPFS meets NICE's criterion for surrogate validity. The results of the surrogacy analysis as applied to the ARANOTE trial are presented in the CS as a sensitivity analysis (CS section B.3.9.4.1), and are summarised in this report in section 4.4.3,

4.4 Results from the ITC

Below is a brief summary and EAG interpretation of the results of the ITC reported in the CS, for each outcome measure in turn. A summary tabulation of the ITC results can be found in Appendix 3.

4.4.1 rPFS

The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a HR of [redacted] In this analysis an HR of [redacted] favours apalutamide + ADT, and the credible interval [redacted] confirming no statistically significant difference between the two treatments.

Apalutamide + ADT had the best mean rank of being effective (1.247 (95% CrI 0.998, 2.000)), with darolutamide + ADT having the second-best mean rank [redacted] and placebo + ADT ranked third [redacted]. The mean rank credible intervals for darolutamide + ADT and apalutamide + ADT overlapped.

The SUCRA values were 88%, [redacted] and [redacted] for the three treatments, respectively. The intervention with the highest SUCRA value would be regarded as the best, and in this case apalutamide + ADT is the best performing treatment of the three. The results of the random effects model were similar to the fixed-effect model, with wider credible intervals as would be expected.

4.4.2 OS

The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a HR of [redacted] In this analysis an HR of [redacted] favours apalutamide + ADT, and the credible interval [redacted] confirming no statistically significant difference between the two treatments. Apalutamide + ADT had the best mean rank of being effective (1.169 (95% CrI 0.998, 2.000)) with darolutamide + ADT having the second-best mean rank ([redacted] with placebo + ADT ranked third [redacted]. The mean rank credible intervals for darolutamide + ADT and apalutamide + ADT overlapped. The SUCRA values were 92%, [redacted] and [redacted] for the three treatments, respectively indicating that apalutamide + ADT is the best performing treatment of the three. The results of the random effects model were similar, with wider credible intervals as would be expected.

4.4.3 Surrogate OS sensitivity analysis

As mentioned earlier (section 4.3.2) the company did a sensitivity analysis using surrogate OS estimates from ARANOTE and reported OS estimates for TITAN (CS Section B.3.9.4.1). The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a HR of [redacted] In this analysis an HR of [redacted] favours apalutamide + ADT, but the credible interval [redacted] confirming no statistically significant difference between the two treatments.

The SUCRA and mean rank are [redacted] for apalutamide + ADT and darolutamide + ADT, with almost overlapping 95% CrIs on the latter (SUCRA 74% versus [redacted], respectively; mean rank 1.516 (1.000, 2.000) versus [redacted]

The surrogate OS sensitivity analysis was repeated using a random effects model (CS Appendix D, Section D1.7.6). Results were similar to the fixed-effect model, but with wider credible intervals, as predicted.

4.4.4 Time to deterioration in FACT-P

The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a HR of [redacted], meaning that people treated with darolutamide + ADT are [redacted] more likely to experience a longer time to deterioration in FACT-P total score than people

treated with apalutamide + ADT. In this analysis an HR of [redacted] favours darolutamide + ADT, and the credible interval [redacted] confirming a marginal statistically significant difference between darolutamide + ADT and apalutamide + ADT.

Darolutamide + ADT had the best mean rank of being effective ([redacted]) with placebo + ADT having the second-best mean rank ([redacted] and apalutamide + ADT ranked third (2.565 (95% CrI 1.997, 3.004). The mean rank credible intervals for darolutamide + ADT and apalutamide + ADT did not overlap. The SUCRA values were [redacted], [redacted] and 22% for the three treatments respectively, indicating that darolutamide + ADT is the best performing treatment of the three. The results of the random effects model were similar to the fixed-effect model, with wider credible intervals, as would be expected. The point estimate results of the random effects model were similar to the fixed-effect model, but the wider credible intervals meant that the credible interval for the HR [redacted]. Therefore, the difference between darolutamide + ADT and placebo + ADT was no longer statistically significant and there is greater uncertainty due to considerably wider credible intervals (darolutamide + ADT vs apalutamide + ADT random effects HR [redacted] ([redacted].

4.4.5 Grade 3-5 adverse events

The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a risk ratio (RR) of [redacted], meaning that people treated with darolutamide + ADT had a [redacted] reduction in the risk of experiencing Grade 3-5 adverse events than those treated with apalutamide + ADT. In this analysis a RR of [redacted] favours darolutamide + ADT, but the credible interval [redacted] confirming no statistically significant difference between the two treatments.

Placebo + ADT had the best mean rank for reduced risk of adverse events ([redacted]) with darolutamide + ADT having the second-best mean rank ([redacted] and apalutamide + ADT ranked third (2.842 (95% CrI 1.998, 3.002). The mean rank credible intervals for darolutamide + ADT and apalutamide + ADT overlapped. The SUCRA values were [redacted], [redacted] and 8% for the three treatments, respectively indicating that placebo + ADT is the best performing treatment of the three. The results for placebo + ADT and darolutamide + ADT are remarkably [redacted]. The results of the random effects model were similar to the fixed-effect model, with wider credible intervals, as would be expected.

4.4.6 Discontinuation due to adverse events

The fixed-effect indirect comparison of darolutamide + ADT and apalutamide + ADT resulted in a RR of [redacted] meaning that people treated with darolutamide + ADT had a [redacted] reduction in the risk of discontinuing treatment due to adverse events than people treated with apalutamide + ADT. In this analysis a RR of [redacted] favours darolutamide + ADT, and the credible interval [redacted] confirming a statistically significant difference between the two treatments.

Darolutamide + ADT had the best mean rank for lower risk of discontinuation due to adverse events ([redacted]) with placebo + ADT having the second-best mean rank ([redacted] and apalutamide + ADT ranked third (2.998 (95% CrI 2.998, 3.002). The mean rank credible intervals for darolutamide + ADT and apalutamide + ADT did not overlap. The SUCRA values were [redacted], [redacted] and 0% for the three treatments, respectively indicating that darolutamide + ADT is the best performing treatment of the three. The point estimate results of the random effects model were similar to the fixed-effect model, but the wider credible intervals meant that the credible interval for the risk ratio [redacted]Therefore, the difference favouring darolutamide + ADT was no longer statistically significant and there is less certainty in effects when a random-effects model is used.

4.4.7 Summary of the results from the ITC

The results from the fixed-effect analysis for efficacy relating to disease progression (rPFS) and survival (OS) show differences in treatment effects in favour of apalutamide + ADT. However, the differences are not statistically significant and should not compromise a cost-comparison analysis.

In contrast, the results of the fixed-effect analysis for quality of life (time to deterioration in FACT-P total score) and adverse events (Grade 3-5 AEs and discontinuation to AEs) show a difference in treatment effect in favour of darolutamide + ADT. The differences in treatment effect are statistically significant for time to deterioration in FACT-P total score and for discontinuation due to AEs. The selection of outcomes that illustrate the known tolerability of darolutamide may bias results in favour of treatment with darolutamide + ADT but interpreting the results as evidence of similarity for a cost-comparison analysis they highlight a small difference between the treatment effects of the intervention and comparator.

Results from the random effects analysis are similar to the fixed-effect models, in terms of point estimates, but none are statistically significant due to wider credible intervals estimated according to random effects assumptions.

Overall, the ITC results suggest no differences between darolutamide + ADT compared to apalutamide + ADT except for a couple of instances when statistically significant differences favoured darolutamide and ADT. This is consistent with the requirement for health technologies to provide similar or greater health benefits to existing recommended treatments. However, a caveat to this that the EAG's observation (discussed earlier in section 4.3.5.3) that, due to differences in the magnitude of patient crossover between the ARANOTE and TITAN trials, the ITC is underestimating the relative effect of apalutamide + ADT. We return to this issue in the following section.

4.4.8 Summary of EAG critique of the ITC methods

The ITC has several strengths but some key limitations which indicate uncertainty in the results and conclusions.

4.4.8.1 Strengths of the ITC

- The ITC is based on a comprehensive SLR which did a systematic search for relevant studies to facilitate an evidence network. The EAG is not aware of any relevant studies which were not identified.
- The two studies included are both pivotal phase III multi-centre, double-blind RCTs – the ARANOTE trial comparing darolutamide + ADT versus placebo + ADT, and the TITAN trial comparing apalutamide + ADT versus placebo + ADT. Both are well-designed trials at low risk of bias during the double-blind phase (but see limitations below).
- A reasonably comprehensive ITC feasibility assessment was undertaken, which gave particular attention to clinical heterogeneity. The trials are generally similar in design, and measurement of outcomes, and patient characteristics (but see below).
- The company's ITC is an NMA using a Bayesian generalised linear model framework to estimate the relative efficacy and safety of darolutamide +ADT versus apalutamide + ADT, based on NICE DSU TSD 2. This is a standard approach to NMA used in NICE technology appraisals. The model appears to have been implemented appropriately.
- The methods used in the NMA are well reported. The process of random effects versus fixed-effect model fitting and selection is transparent and well justified, and results from both models are available for all outcomes and show consistency in conclusions.

4.4.8.2 Limitations of the ITC

4.4.8.2.1 Effects of patient crossover

The EAG considers one of the main limitations of the ITC is that in both trials final OS is based on follow-up data collected in the open-label trial phase during which placebo patients crossed over to the experimental treatment.

- In both trials patients were analysed using an ITT approach, resulting in confounding in the placebo group estimates from inclusion of crossed-over patients receiving darolutamide / apalutamide. This potentially underestimates the relative effects of darolutamide/apalutamide + ADT versus placebo + ADT in the respective trials.
- Selected crossover adjusted OS estimates from the trials are reported but have not been included in the NMA. There is no discussion in the CS of the implications of using ITT or crossover adjusted effect estimates as input parameters. The EAG considers it appropriate in this current appraisal to use the unadjusted estimates in the NMA as a base case, since the direction of bias in in both is expected to be the same (i.e. underestimating the effect of the experimental treatment).

However, the magnitude of the bias appears to be larger in the TITAN trial, in which a higher percentage of placebo group patients crossed over to apalutamide and median follow up was longer. The ITT and crossover adjusted OS estimates are similar in the ARANOTE trial (see

- Table 6) but in the TITAN trial the crossover adjusted OS HR was noticeably lower than the ITT HR, illustrating underestimation in in the relative effects of apalutamide + ADT versus placebo + ADT in the trial.
- Using the ITT based rather than crossover adjusted OS estimates in the ITC potentially underestimates the relative efficacy of apalutamide compared to darolutamide. The crossover adjusted HR would likely result in a larger reduction in death favouring apalutamide, potentially shifting the upper bound of the current OS HR credible interval to less than 1, indicating a statistically significant difference.
- However, the EAG urges caution in the interpretation of the crossover adjusted estimates, because only a limited selection of adjustment methods were reported (two for the ARANOTE trial, and only one for the TITAN trial). It is unclear how robust the OS estimates are when adjusted according to methods using alternative assumptions.
- If crossover-adjusted OS estimates were used in the ITC this would be best viewed as an exploratory sensitivity analysis. The EAG considers the ITT based OS

estimates are more appropriate for the base case, bearing in mind the uncertainty outlined above regarding the potential underestimation of the relative effects of apalutamide + ADT in TITAN.

-

The results of the ITC for the rPFS outcome are not affected by crossover since they reflect only the double-blind phase of the trials. The relative effects of darolutamide + ADT versus apalutamide + ADT on rPFS can therefore be regarded as more certain.

4.4.8.2.2 Inferring similarity of effects

The CS states that the results of the ITC for rPFS show “no evidence of a difference” (CS page 70) between darolutamide + ADT and apalutamide + ADT, the implication therefore being that they demonstrate “similar efficacy” (CS page 5), hence supporting the company’s case for a cost comparison appraisal. The EAG agrees that most of the ITC analyses do not show a statistically significant difference between darolutamide and apalutamide, but this does not necessarily imply they are similar in effects. The most appropriate method of establishing similarity would be from an equivalence or non-inferiority trial directly comparing darolutamide + ADT versus apalutamide + ADT. Such a trial would require a large enough sample of patients to demonstrate equivalence or non-inferiority within pre-defined effect margins. The sample sizes of the respective ARANOTE and TITAN trials were set for the purpose of confirming superiority over placebo and are not necessarily sufficient for detecting equivalence / non-inferiority. In the absence of such a trial the ITC is nonetheless informative though its limitations should be taken into account.

5 COST COMPARISON MODEL

5.1 Model structure and assumptions

The company conducted a simple cost comparison analysis in Microsoft Excel comparing the drug acquisition costs of darolutamide + ADT with those of apalutamide + ADT in the treatment of adults with mHSPC. Patient outcomes over time were not modelled through a health economic model (such as adopting a Markov approach or a partitional survival modelling approach). The company further justified their simplified approach in their response to EAG clarification question B1, maintaining that the only difference between the intervention and comparator arms is the drug acquisition costs of darolutamide and apalutamide.

In their cost comparison analysis, the company made the following assumptions:

- No differences in the resource use to administer darolutamide and apalutamide
- No differences in treatment monitoring and managing adverse events between darolutamide and apalutamide

EAG comment on the model structure and assumptions

We view that it would be appropriate to provide a cost-comparison model that incorporated: i) clinical efficacy (e.g. survival estimates including PFS and OS); ii) costs (including, drug acquisition, drug administration, resource use, subsequent treatments, adverse events); iii) safety outcomes (e.g. adverse events); and iv) mortality. If the model included the above parameters it would enable the EAG to perform a robust verification of the model assumptions. Currently, the EAG are unable to test the impact of varying the parameter inputs such as resource use, subsequent treatments, and adverse events, on the overall cost-comparison results. However, based on the clinical evidence (discussed in section 4) and our expert's clinician's opinion, we view that darolutamide and apalutamide are likely to have similar effectiveness and resource use, as we will discuss in section 5.2.

5.1.1 Model features

The cost comparison analysis included the following features:

- **Population:** Adult men with mHSPC who are unsuitable for chemotherapy. This is narrower than the defined population in the NICE scope; the company restricts the patient population to those who cannot have docetaxel. The mean age of the modelled cohort is 69 years, based on the ARANOTE[17] trial.

- **Intervention:** Darolutamide with ADT. This aligns with the NICE scope.
- **Comparator:** Apalutamide with ADT. The EAG considers this to be acceptable. NICE TA741 recommends apalutamide for people with mHSPC who are unsuitable for chemotherapy. The company appear not to have considered enzalutamide with ADT as a comparator. They did not provide an explicit statement about why it was not considered, rather, they justified their chosen comparator, apalutamide, on the basis that it is recommended at the same point in the treatment pathway as is anticipated for darolutamide. The clinical experts advising the EAG viewed that enzalutamide and apalutamide are both commonly used in clinical practice.
- **Perspective**: The company state that the perspective for costing is that of the UK NHS and PSS. An NHS and PSS perspective is appropriate for the NICE Reference Case.
- **Time horizon**: effectively lifetime - 25 years (maximum age 100 years)
- **Cycle length**: 28-day cycle length
- **Half cycle correction:** Not applied
- **Discounting**: 3.5% per annum applied to drug acquisition.
- **Mortality**: The company assumes equal mortality across treatment arms. However, mortality is excluded from the analysis. This may be a reasonable assumption given that OS is similar between the two treatment arms (see section **Error! Reference source not found.**).

5.2 Model parameters

5.2.1 Time on treatment

The company uses time on treatment (ToT) in estimating the drug acquisition costs. Within the economic model, the proportion of patients on treatment in each cycle are multiplied by the drug acquisition costs. The CS states that the proportion of patients in the intervention arm were informed by the ARANOTE[17] trial and the same ToT was applied to the comparator.. Kaplan Meier curves for ToT for darolutamide + ADT along with the standard parametric models including exponential, Weibull, log-normal, log-logistic, Gompertz, Gamma and generalised gamma are presented in CS Figure 12; the corresponding Akaike Information Criterion (AIC) values in CS Table 31. The company applied the log-logistic curve in their base case and explored the use of gamma and generalised gamma in the scenario analyses. Furthermore, they conducted a scenario analysis where drug costs were adjusted based on radiographic PFS (rPFS), instead of ToT, assuming all patients would be treated up to progression or death, whichever occurred first. For this scenario, they applied

the log-normal distribution to extrapolate the rPFS Kaplan Meier curve for darolutamide + ADT, obtained from the ARANOTE trial.

EAG comment

Overall, we view the company's approach is reasonable. Estimating drug acquisition costs based on ToT is a conservative assumption, compared to that based on rPFS.

5.2.2 Drug acquisition costs

As stated in the previous section, ToT data from the ARANOTE trial was applied to the darolutamide and apalutamide arms to estimate the drug acquisition costs. The company justified their approach citing that there was an absence of ToT data for apalutamide. Information on dosing regimens, dose intensity and unit costs for the treatment arms are in CS Table 33. They used the list prices of the drugs from the British National Formulary (BNF)[18 19] and applied a confidential price discount of [redacted] on the price of darolutamide.

For ADT, the company applied the list prices obtained from the BNF. In response to clarification question B4, the company provided the information on dosing regimens, dose intensity and unit costs for the ADTs in Table 5 of the clarification response document. However, the EAG noted an inconsistency in the proportions of the constituent ADT treatments as reported in the CS and the economic model (shown in below in Table 7). In their response to clarification question B3, the company acknowledged the inconsistency and clarified that the values reported in the CS (which are based on NICE TA903[20] 'Darolutamide in combination with docetaxel and ADT') are appropriate. They corrected the values in their revised model submitted as part of the clarification response. This correction reduced the per capita cycle costs of ADT acquisition costs from £119.30 to £68.09 and administration costs from £435.98 to £238.96, respectively. However, the change does not impact the overall results of the cost comparison analysis as patients in both the treatment arms are assumed to have the same duration of treatment with ADT.

Table 7 Distribution of ADT treatments included in the model

Treatment	Administration route	Mix Proportion	
		CS	Economic Model
Degarelix	SC injection	Not reported	12.6%
Leuprorelin	SC injection	30.0%	54.0%
Goserelin	SC injection	30.0%	31.9%
Triptorelin	Oral	40.0%	1.5%
Buserelin	Oral	Not reported	0%

Source: CS model and CS B.4.2.3
ADT, Androgen deprivation therapy; SC, subcutaneous

The company assumed 100% relative dose intensity for darolutamide + ADT and apalutamide + ADT.

5.2.3 Drug administration costs

Drug administration costs were excluded in the model analysis as the company assumed the same rate of disease progression and ToT between darolutamide and apalutamide. Both darolutamide and apalutamide have oral administration with a daily dosing schedule. The ADT constituent treatments are the same for both the treatment arms.

5.2.4 Healthcare resource use and associated costs

Healthcare resource use (HCRU) is excluded from the cost-comparison analysis. The company state that darolutamide offers several benefits over apalutamide, leading to less resource use (such as, consultation with GPs, oncologists, pharmacists) and easier monitoring and patient management. The company argues that patients receiving apalutamide would require thyroid function tests as part of treatment monitoring as well as additional steps to ensure patient safety due to apalutamide's higher number of drug-drug-interactions. However, the company state that they have adopted a conservative approach and assumed comparable HCRU between the two treatments arms.

5.2.5 Adverse reaction unit costs and resource use

The economic model excluded any adverse event related costs and resource use. The CS presented the differences in adverse events associated with darolutamide + ADT in the ARANOTE trial compared to those of apalutamide + ADT in the TITAN trial in CS Table 34. Based on the ITC findings, the summary of the findings from the ARANOTE and TITAN trials, and their expert clinical opinions, the company argue that the safety profile of darolutamide is likely to be similar to, or better than, that of apalutamide.

EAG comment on model parameters

The model parameters are programmed correctly in the Excel spreadsheet. Expert clinical advice to the EAG supports the assumption that darolutamide is likely to require less health care resource and incur in fewer costs compared to apalutamide. However, we could not assess the impact of varying assumptions about resource use because this functionality is not included in the company model.

5.3 EAG model checks

The company did not mention model validation in their submission. The EAG checks of the company's cost comparison model included:

comparing all parameter values against the CS and the cited source documents;

checking the calculations in the MS Excel spreadsheet, and

double programming parts of the model, i.e., constructing a duplicate model version to check that it produced the same results.

We noticed that:

- the half-cycle modelling was not implemented.
- the administration costs were declared ("Treatment costs" sheet, cells M34 to R42), but not considered in the cost-comparison model. However, there is no effect in the model results as the intervention and comparator were assumed to have the same ADT regimen and ToT / rPFS curves.

We were able to reproduce the original model results (base case and scenarios). We confirm that the evidence sources and the values applied in the cost-comparison model are consistent with their sources, except for:

- A minor difference in the mean age (69 years old in the CS, 69.67 years old in the economic model, and 69.6 years old in ARANOTE CSR[17]). In response to clarification question C2, the company confirmed that the mean age is 69 years old and amended the model.
- There is a difference in the proportions of the constituent ADT treatments between the CS and the cost comparison model (see Table 7). In response to clarification question B3, the company stated that the correct proportions are presented in the CS and amended them in the model. The updated acquisition cost for the ADTs per cycle is £68.09, and the updated administration cost for the ADTs is £238.96.
- The company's base case results remained the same, as the discrepancies above did not affect the total incremental cost.

EAG comment on model checking and validation:

The cost-comparison model is generally well implemented. However, we spotted minor discrepancies between the CS and the original cost comparison model which the company duly corrected. The EAG has implemented the half-cycle correction in the corrected cost comparison model and presented the results in section 6.3 below.

6 COMPANY AND EAG COST COMPARISON RESULTS

6.1 Company cost comparison results

The total cost is based on the drug acquisition costs of the intervention, comparator and ADT medications and is shown in CS section B 4.3. The company base case results with the PAS discount price for darolutamide ([redacted]) and the list price for apalutamide and ADTs are in CS Table 35. The company's base case results with the list price for the intervention, comparator, and ADTs are in CS Table 36.

The results in CS Table 35 suggest that darolutamide + ADT [redacted] relative to apalutamide + ADT with the incremental cost of [redacted]. The company's corrections mentioned in section 5.3 did not affect the company's base case incremental costs but affected the ADT acquisition cost. Table 8 below shows the company base case updated results using the revised cost comparison model provided by the company with the clarification responses. The EAG notes that these analyses include the PAS price only for darolutamide, and list prices for apalutamide and the ADTs. We report results using the PAS discount prices for all treatments (where applicable) in a separate confidential addendum to this report.

Table 8 Company's base case updated results: PAS price for darolutamide and list price for apalutamide and ADT medications

	Darolutamide + ADT	Apalutamide + ADT	Difference
Drug acquisition	[redacted]	£146,218	[redacted]
Total cost	[redacted]	£146,218	[redacted]

Source: Revised cost comparison model
ADT, Androgen deprivation therapy; PAS, Patient access scheme

6.2 Company sensitivity and scenario analyses

The company did not provide deterministic and probabilistic sensitivity analyses. The EAG agrees that the cost comparison model parameters described in section 5.1 are more suited to scenario analysis as a method to explore uncertainty. The company scenario analyses are described in section CS section B.4.4. The scenario analyses results for darolutamide + ADT vs. apalutamide + ADT are in CS Table 37 (PAS price for darolutamide and list prices for apalutamide and ADTs) and CS Table 38 (list prices for darolutamide, apalutamide, and ADTs).

Considering the PAS price for darolutamide (CS Table 37), all scenarios [redacted] with the incremental cost varying from [redacted] The incremental cost remained the same when we ran the scenarios using the revised cost comparison model. Results with confidential price discounts for apalutamide and ADTs are reported in a confidential addendum to this report.

6.3 EAG's cost comparison results

Table 9 below shows the results with the EAG's correction to the company's cost comparison model, including the half-cycle correction mentioned in section 5.3. The incremental total cost varied from [redacted] to [redacted].

Table 9 EAG correction to the company base case updated results: PAS price for darolutamide and list price for apalutamide and ADT medications

	Darolutamide + ADT	Apalutamide + ADT	Difference
Drug acquisition	[redacted]	£145,022	[redacted]
Total cost	[redacted]	£145,022	[redacted]

Source: EAG corrected cost comparison model
PAS, Patient access scheme; ADT, Androgen deprivation therapy

6.4 EAG's scenarios

Table 10 below shows the company scenario results using the EAG corrected cost comparison model and EAG scenarios. For the EAG scenarios, we noticed that:

- The company's base case uses the log-logistic distribution (based on the lowest AIC model fit value) to model the ToT curve. Changing the ToT distribution curve varied the incremental total cost from [redacted].
- The company's scenario analysis, in which the drug cost adjustment was based on rPFS, has an incremental cost of [redacted] (scenario 6: log-normal, lowest AIC). Changing the distribution curve assigned to rPFS varied the incremental total cost from [redacted].
- Changing the discount rate from 3.5% to 1.5% resulted in an incremental cost of [redacted] (scenario 7).

The following scenarios did not affect the incremental cost result in the company's base case: varying mean age, varying the proportion of the ADTs, and the inclusion of the administration cost of the ADTs.

Table 10 Company and EAG scenarios: PAS price for darolutamide and list price for apalutamide and ADT medications

Base case	ID	Scenario	Darolutamide + ADT	Apalutamide + ADT	Incremental cost
EAG corrected base case result			[redacted]	£145,022	[redacted]
Company scenarios					
Time horizon: 25 years	1	10 years	[redacted]	£123,728	[redacted]
	2	15 years	[redacted]	£135,182	[redacted]
Discounting: 3.5%	3	No discounting	[redacted]	£173,565	[redacted]
Alternative ToT extrapolations: log-logistic	4	Gamma	[redacted]	£111,579	[redacted]
	5	Generalised gamma	[redacted]	£124,805	[redacted]
Drug cost adjustments based on ToT	6	Based on rPFS: log-normal	[redacted]	£204,363	[redacted]
EAG scenario					
Discount rate: 3.5%	7	1.5%	[redacted]	£159,810	[redacted]
Apply ToT adjustment: log-logistic extrapolation curve:	8	Exponential	[redacted]	£125,833	[redacted]
	9	Weibull	[redacted]	£144,128	[redacted]
	10	Log-normal	[redacted]	£160,133	[redacted]
	11	Gompertz	[redacted]	£104,064	[redacted]
Apply ToT adjustment: log-logistic	12	Apply rPFS: Exponential	[redacted]	£174,754	[redacted]
	13	Apply rPFS: Weibull	[redacted]	£199,916	[redacted]
	14	Apply rPFS: Log-logistic	[redacted]	£187,425	[redacted]
	15	Apply rPFS: Gompertz	[redacted]	£127,441	[redacted]
	16	Apply rPFS: Gamma	[redacted]	£144,465	[redacted]
	17	Apply rPFS: Generalised gamma	[redacted]	£249,018	[redacted]

Source: EAG corrected cost comparison model
ADT, Androgen deprivation therapy; EAG, External Assessment Group; PAS, Patient Access Scheme; rPFS, radiographic progression free survival; ToT, Time on Treatment

6.5 EAG's conclusion on the cost comparison

The company provided a cost comparison model that estimated only the difference in the drug acquisition costs between the darolutamide + ADT and apalutamide + ADT. Although this cost comparison model is aligned with the guideline ("User guide for the cost comparison company evidence submission template (PMG32)")[21], the EAG are unable to test the impact of varying the parameter inputs such as resource use, subsequent treatments, and adverse events, on the overall cost-comparison results.

The company's results suggest that, compared with apalutamide + ADT, darolutamide + ADT is associated with lifetime cost savings for patients with mHSPC when using the discounted PAS price for darolutamide and list price for apalutamide and ADTs (leuprorelin, goserelin and triptorelin). The EAG corrected the company's cost comparison model (see section 5.3), with marginal impact on the total cost (incremental total cost varied from [redacted] to [redacted]).

We report results for the company's and EAG's analysis using all available NHS price discounts for apalutamide and ADTs in a confidential addendum to this report.

7 EQUALITIES AND INNOVATION

The CS notes the presence of a small but significant equality gap in the mHSPC treatment pathway affecting people with a history of experiencing seizures or other predisposing factors. The currently available ARTAs are contraindicated in this group of people reducing their available treatment options to ADT monotherapy, which is considered sub-optimal by today's standards. The CS highlights that darolutamide has the potential to address this inequality as it is not contraindicated in such patients. As mentioned earlier, the EAG's clinical expert recognised that darolutamide can be used in people with seizures and other central nervous system disorders, although she also noted that such patients are rarely seen in clinical practice.

The disproportional impact of prostate cancer on particular population groups, notably Black males, older/elderly people and people with comorbidities is discussed in the CS. It is noted that treatment intensification (which refers to strategies to combine existing treatments with additional therapies, such as chemotherapy, to improve outcomes) decreases in these groups, though it is not explicitly stated why (e.g. contraindications/intolerance to adverse events/poor access to health care). The company point to the need for additional treatments for these groups, with darolutamide presumably filling this gap.

The CS does not explicitly discuss innovation in relation to darolutamide. By its nature, the cost comparison approach implies that the health technology under appraisal is not the first treatment of its kind. Rather, it shares similarities with current established therapies. The EAG suggests that, although darolutamide is not the first second-generation ARTA for treatment of mHSPC, its distinct chemical structure differentiates it from apalutamide and enzalutamide. Consequently, darolutamide is associated with reduced blood-brain barrier penetration and low central nervous system side effects, making it suitable for use in patients contraindicated to current treatments. This can be regarded as an innovative feature of darolutamide which adds value over current treatment options.

8 EAG COMMENTARY ON THE ROBUSTNESS OF EVIDENCE SUBMITTED BY THE COMPANY

The EAG considers the evidence submitted by the company appropriately supports a cost comparison appraisal. However, it is important to acknowledge the uncertainties discussed earlier.

The assumption that darolutamide and apalutamide are similar in efficacy and safety rests upon the company's ITC (NMA). The ITC uses standard methods and assumptions and for the most part is transparently reported. However, the ITC is limited by the sparse available data available on the relative efficacy and safety of darolutamide compared to apalutamide. ARANOTE and TITAN are well-designed multi-centre double blind RCTs but due to their designs there is potential for confounding in the final survival estimates, and for this to be carried through into the ITC. The ITC results showed no statistically significant difference between darolutamide and apalutamide across the various analyses undertaken, except for a couple of outcomes (one of which is adverse events), indicating the superiority of darolutamide over apalutamide. We have raised the possibility of darolutamide being found inferior to apalutamide in terms of OS but this remains to be tested. In the meantime the uncertainty remains.

Despite the above concerns the EAG is inclined to adopt a pragmatic view and suggest that, in the absence of further evidence, the similarities in chemical composition and mechanism of action shared by darolutamide and the other ARTAs, endorsed by expert clinical opinion, provides a sufficient basis upon which to assume general similarity in efficacy and safety, and thus support for a cost comparison appraisal.

A further issue is that it's not possible for the EAG to test the impact on survival, and of varying the cost comparison model parameter inputs, such as resource use, subsequent treatments, and adverse events. This is because the structure of the company's model does not cater for these analyses. However, based on clinical effectiveness evidence and expert clinical advice the EAG understands that darolutamide + ADT is likely to have similar efficacy, similar use of resources and costs (see sections 5.1 and 5.2) to apalutamide + ADT. Nonetheless, the EAG would prefer to be able to independently test these assumptions.

9 REFERENCES

1. NHS England. Clinical Commissioning Policy Statement: Docetaxel in combination with androgen deprivation therapy for the treatment of hormone naïve metastatic prostate cancer (B15/PS/a). Updated: 20 January 2016 [Available from: https://www.england.nhs.uk/wp-content/uploads/2016/01/b15psa-docetaxel-policy-statement.pdf?UNLID=106952065020223157562 4 accessed 22 October 2024.

2. EMA. Summary of Product Characteristics: darolutamide (NUBEQA) 2024 [updated 20-Dec-2024. Available from: https://www.medicines.org.uk/emc/product/11324/smpc/print.

3. NICE. Enzalutamide for treating hormone-sensitive metastatic prostate cancer. Technology appraisal guidance TA712, 07 July 2021.

4. BNF. British National Formulary: darolutamide drug interactions 2025 [Available from: https://bnf.nice.org.uk/interactions/darolutamide/ accessed 11 April 2025.

5. BNF. British National Formulary: enzalutamide drug interactions 2025 [Available from: https://bnf.nice.org.uk/interactions/enzalutamide/ accessed 11 April 2025.

6. BNF. British National Formulary: apalutamide drug interactions 2025 [Available from: https://bnf.nice.org.uk/interactions/apalutamide/ accessed 11 April 2025.

7. Saad F, Vjaters E, Shore M, Olmos D, Xing N, Pereira de Santana Gomes AJ, et al. Darolutamide in combination with androgen-deprivation therapy in patients with metastatic hormone-sensitive prostate cancer from the phase III ARANOTE trial. ESMO September 16, 2024. *J Clin Oncol* 2024;42(36):4271-81. doi: https://doi.org/10.1200/JCO-24-01798

8. Chi KN, Agarwal N, Bjartell A, Chung BH, Pereira de Santana Gomes AJ, Given R, et al. Apalutamide for metastatic, castration-sensitive prostate cancer. *N Engl J Med* 2019;381(1):13-24. doi: 10.1056/NEJMoa1903307 [published Online First: 20190531]

9. Bayer AG. Final Overall Survival Results Summary for Study 21140 (ARANOTE): A randomized, double-blind, placebo-controlled Phase 3 study of darolutamide in addition to androgen deprivation therapy (ADT) versus placebo plus ADT in men with metastatic hormone-sensitive prostate cancer (mHSPC). Data cut-off: 16 Jan 2025., 2025.

10. Esper P, Mo F, Chodak G, Sinner M, Cella D, Pienta KJ. Measuring quality of life in men with prostate cancer using the Functional Assessment of Cancer Therapy-prostate instrument. *Urology* 1997;50(6):920-28. doi: https://doi.org/10.1016/S0090-4295(97)00459-7

11. Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating Clinically Meaningful Changes for the Functional Assessment of Cancer Therapy—Prostate: Results from a Clinical Trial of Patients with Metastatic Hormone-Refractory Prostate Cancer. *Value Health* 2009;12(1):124-29. doi: https://doi.org/10.1111/j.1524-4733.2008.00409.x

12. Chi KN, Chowdhury S, Bjartell A, Chung BH, Pereira de Santana Gomes AJ, Given R, et al. Apalutamide in patients with metastatic castration-sensitive prostate cancer: final survival analysis of the randomized, double-blind, phase III TITAN study. *J Clin Oncol* 2021;39(20):2294-303. doi: 10.1200/JCO.20.03488 [published Online First: 20210429]

13. NICE. NICE Technology Appraisal TA741: Apalutamide with androgen deprivation therapy for treating hormone-sensitive metastatic prostate cancer – Technology appraisal guidance. October 28, 2021. 2021

14. Cooper K, Tikhonova I, Woods L, Hazel L, Souto Ribeiro I, Picot J, et al. Apalutamide for treating prostate cancer - A Single Technology Appraisal: Southampton Health Technology Assessments Centre, University of Southampton, 2020.

15. Dias SWNSA. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. Report by the Decision Support Unit 2011 (last updated September 2016), 2011.

16. Bujkiewicz S, Achana F, Papanikos T, Riley RD, Abrams KR. NICE DSU Technical Support Document 20: multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints. 2019. *Available from: https://wwwsheffieldacuk/nice-dsu/tsds/full-list* 2019

17. Bayer. Clinical Study Report: Phase 3 study of darolutamide in addition to ADT versus placebo plus ADT in men with mHSPC (ARANOTE). 2024

18. BNF. Darolutamide. Medicinal forms [Available from: https://bnf.nice.org.uk/drugs/darolutamide/medicinal-forms/ accessed 27/01/2025 2025.

19. BNF. Apalutamide. Medicinal forms [Available from: https://bnf.nice.org.uk/drugs/apalutamide/medicinal-forms/ accessed 27/01/2025 2025.

20. NICE. Darolutamide with androgen deprivation therapy and docetaxel for treating hormone-sensitive metastatic prostate cancer. Technology appraisal guidance [TA903] 2023 [Available from: https://www.nice.org.uk/guidance/ta903.

21. NICE. User guide for the cost comparison company evidence submission template (PMG32), 2017:41.

22. McKay RR, E. RA, A. PM, R. GJ, S. SS, Natasha L, et al. Darolutamide plus androgen-deprivation therapy: propensity score matching of ARASEC and historic clinical trial patients. *Future Oncology* 2025;21(11):1365-75. doi: 10.1080/14796694.2025.2482360

10 APPENDICES

Appendix 1 EAG appraisal of the clinical SLR methods

Table 11 Summary of the EAG appraisal of the clinical SLR methods

Systematic review components and processes	EAG response	EAG comments
Was the review question clearly defined using the PICOD framework or an alternative?	Yes	Criteria for inclusion are organised according to a PICOD framework (CS Appendix Table 14). The criteria include all treatments for mHSPC, not only those relevant to this submission (CS Appendix D.1.1.2).
Were appropriate sources of literature searched?	Yes	MEDLINE (including In-process records), Embase and Cochrane databases were searched, additionally EBM Reviews, relevant conferences, and bibliographies of relevant systematic reviews and meta-analyses published in the last 6 years (CS Appendix D.1.1.1).
What time period did the searches span and was this appropriate?	Yes	An original search and three update searches were carried out covering from database inception to 24 October 2024 (CS Appendix D.1.1.1). There were no gaps in coverage, and they are reasonably up to date.
Were appropriate search terms used and combined correctly?	Yes	The search terms were comprehensive and relevant; they were combined appropriately in the databases (CS Appendix Tables 1 to 13).
Were inclusion and exclusion criteria specified? If so, were these criteria appropriate and relevant to the decision problem?	Yes	Criteria for inclusion are in CS Appendix Table 14. They are for a global SLR therefore include all treatments used for mHSPC. The excluded studies list consistently excluded ARANOTE, the pivotal trial, due to study design which raised concerns as to whether all relevant studies were identified. The EAG performed targeted searches in MEDLINE and ClinicalTrials.gov and identified the company ARASEC trial[22] which compares darolutamide + ADT compared to a matched historical control arm of ADT alone (derived from the CHAARTED RCT) for treating men with mHSPC in the United States. The trial publication was published after the last search date, but we identified it via ClinicalTrials.gov (NCT05059236). This study could potentially contribute evidence to an ITC

Systematic review components and processes	EAG response	EAG comments
		however, the company confirmed that the completion date for the primary outcome is not expected until Q2 2025, with data available in Q4 2024 9clarification response A16). Therefore, we are now confident that all studies with results relevant to this appraisal were identified.
Were study selection criteria applied by two or more reviewers independently?	Yes	Studies were screened by two reviewers independently with any discrepancies resolved by a third reviewer (CS Appendix D.1.1.2).
Was data extraction performed by two or more reviewers independently?	Unclear	The procedure for conducting data extraction is not reported in either CS Appendix D or CS section B.3.
Was a risk of bias assessment or a quality assessment of the included studies undertaken? If so, which tool was used?	Yes	The company assessed the ARANOTE trial using the NICE checklist for RCTs (CS section B.3.5).
Was risk of bias assessment (or other study quality assessment) conducted by two or more reviewers independently?	Unclear	The number of reviewers conducting the quality assessment is not reported in either CS Appendix D or CS section B.3.5.
Is sufficient detail on the individual studies presented?	Yes	The company provided the relevant clinical study reports for ARANOTE and all relevant published papers with the main submission and the study SAP and protocol for ARANOTE with the clarification response.
If statistical evidence synthesis (e.g. pairwise meta-analysis, ITC, NMA) was undertaken, were appropriate methods used?	Yes	An NMA carried out to compare effectiveness and safety of darolutamide + ADT with apalutamide + ADT. The NMA is discussed in section 4.3 of this report.

Abbreviations: ADT, androgen deprivation therapy; EBM, Evidence Based Medicine Reviews database; ITC, indirect treatment comparison; mHSPC, metastatic hormone-sensitive prostate cancer; NMA, network meta-analysis; PICOD, Population Intervention Comparator Outcomes Design-of-study framework; RTCs, randomised controlled trials; SAP, statistical analysis plan; SLR, systematic literature review.

Appendix 2 Risk of bias assessment for ARANOTE

Table 12 Company and EAG risk of bias assessment for the ARANOTE trial

NICE checklist criteria	Company assessment	EAG assessment
Was randomisation carried out appropriately?	Yes. Randomisation was appropriate and carried out centrally using an Interactive Web Response System (IWRS) system.	Yes, agree. Low risk of bias
Was the concealment of treatment allocation adequate?[a]	Yes. The study was double-blinded such that neither the investigator or study site personnel, the study sponsor or participant knew which drug was being administered. The appearance of darolutamide and placebo were identical, and study drugs were packed in bottles labelled with a unique kit number assigned to the participant via IWRS.	Yes, agree. We assume that the IWRS system ensured that the process of treatment allocation was adequately concealed. **Low risk of bias.**
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes. Patient characteristics were well balanced between the two groups.	Yes, agree. All participant characteristics were similar between groups, including prognostic factors for mHSPC such as disease volume, disease pattern (Gleason score) and presence of visceral metastases. **Low risk of bias.**
Were the care providers, participants and outcome assessors blind to treatment allocation? [b]	Yes, it is a double-blind study.	Yes, agree for outcomes reported at the primary data analysis. The primary analysis is based on assessments made during the double-blind trial period. For rPFS it is also based on BICR. **Low risk of bias for the primary outcome of rPFS.**

NICE checklist criteria	Company assessment	EAG assessment
		Unblinding occurred after the primary completion analysis when participants in the placebo arm were permitted to crossover to darolutamide in the open-label study period. **High risk of bias for outcomes assessed after the primary analysis (i.e. final analyses of OS and adverse events)**
Were there any unexpected imbalances in drop-outs between groups?	No. Authors reported the number of patients and reasons for discontinuation in both treatment groups and these were balanced between groups.	No, agree. However, at the primary completion analysis, more participants in the placebo arm (71.7%) discontinued the study drug than in the darolutamide arm (45.5%). (CS Appendix Figure 2). The biggest cause of discontinuations was disease progression, and was highest in the placebo group, as would be expected in a placebo controlled trial. The remaining reasons for discontinuation are reasonably balanced between arms. **Low risk of bias.**
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No. Authors measured and reported all the outcome as per study primary and secondary endpoints stated in method section.	No, agree. CSR is comprehensive for all outcomes. **Low risk of bias.**
a) Did the analysis include an intention-to-treat analysis? b) If so, was this appropriate?	Yes. This was a FAS analysis for measuring efficacy and mITT for safety outcomes, but no methods were used to account for missing data.	a) Yes, agree. The FAS analysis for efficacy outcomes is analogous to a true ITT analysis. No, disagree for the safety outcomes. The CS refers to

NICE checklist criteria	Company assessment	EAG assessment
c) Were appropriate methods used to account for missing data? [c]		a Safety Analysis Set (SAF) for safety outcomes defined as all participants randomised who took ≥1 dose of study drug. Participants were analysed according to the study drug they received. This is not compatible with an ITT analysis. b) Yes, agree for efficacy outcomes. Yes, agree for safety outcomes. The company's definition of the SAF is similar to that used in other clinical trials and appropriate for attributing adverse events to study drugs. c) No methods were used to account for missing data, with appropriate exceptions reported in CS Table 10. The amount of missing data is not reported so it is unclear whether methods of handling missing data were necessary. **Unclear risk of bias.**

Source: CS Table 11 with added EAG comments; ARANOTE trial publication;[7] ARANOTE Final OS Summary.[9]

[a] The company's justification for answering 'yes' to this question appears to conflate study blinding with allocation concealment. These are two separate procedures in clinical trials, that can lead to different types of bias. See also footnote [b].

[b] The company's justification for their answer to whether allocation concealment was adequate, given earlier in the table (see footnote [a]) is more appropriate as an explanation for their answer to this question on study blinding.

[c] The EAG have split what was a single compound question into three sub-questions (i.e. a, b and c) to enable us to make judgements specific to each sub-question. In contrast, the company's judgements reflect their answer to the original single compound question.

Abbreviations: BICR, blinded independent central review; CSR, clinical study report; FAS, full analysis set; IWRS, Interactive Web Response System; mHSPC, metastatic hormone-sensitive prostate cancer; mITT, modified intention to treat; rPFS, radiographic progression free survival.

Appendix 3 Summary of ITC results

Table 13 Summary of ITC results

Treatment / Outcome	Fixed effects (company base case)				Random effects (company scenario analysis)			
	HR/RR [a] (95% CrI)	SUCRA	Mean rank (95% CrI)	EAG comment	HR/RR [a] (95% CrI)	SUCRA	Mean rank (95% CrI)	EAG comment
rPFS (Hazard Ratio <1 favours apalutamide)								
Apalutamide + ADT	Comparison	0.88	1.247 (0.998, 2.000)	Favours APA; NS	Comparison	0.82	1.370 (1.000, 2.000)	Favours APA; NS
Darolutamide + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Placebo + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Overall survival [b] (Hazard Ratio <1 favours apalutamide)								
Apalutamide + ADT	Comparison	0.92	1.169 (0.998, 2.000)	Favours APA; NS	Comparison	0.86	1.287 (0.999, 2.732)	Favours APA; NS
Darolutamide + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Placebo + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Time to deterioration in FACT-P (Hazard Ratio >1 favours comparator vs apalutamide)								
Darolutamide + ADT	[redacted]	[redacted]	[redacted]	Favours DAR; SS	[redacted]	[redacted]	[redacted]	Favours DAR; NS
Placebo + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Apalutamide + ADT	Comparison	0.22	2.565 (1.997, 3.004)		Comparison	0.35	2.294 (1.000, 3.002)	
Grade 3-5 AEs (Rate Ratio <1 favours comparator vs apalutamide)								
Placebo + ADT	[redacted]	[redacted]	[redacted]	Favours DAR; NS	[redacted]	[redacted]	[redacted]	Favours DAR
Darolutamide + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	

Treatment / Outcome	Fixed effects (company base case)				Random effects (company scenario analysis)			
	HR/RR [a] (95% CrI)	SUCRA	Mean rank (95% CrI)	EAG comment	HR/RR [a] (95% CrI)	SUCRA	Mean rank (95% CrI)	EAG comment
Apalutamide + ADT	Comparison	0.08	2.842 (1.998, 3.002)		Comparison	0.33	2.338 (1.000, 3.002)	over APA; NS
Discontinuation due to AEs (Rate Ratio <1 favours comparator vs apalutamide)								
Darolutamide + ADT	[redacted]	[redacted]	[redacted]	Favours DAR; SS	[redacted]	[redacted]	[redacted]	Favours DAR; NS
Placebo + ADT	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Apalutamide + ADT	Comparison	0.00	2.998 (2.998, 3.002)		Comparison	0.08	2.834 (1.000, 3.002)	

Source: reproduced from CS Tables 16, 18, 20, 22 and 24; CS Appendix D.1.7 Tables 28, 29, 30, 31 and 32.
Abbreviations: ADT, androgen deprivation therapy; AEs, adverse events; APA, apalutamide; CrI, credible interval; DAR, darolutamide; HR, hazard ratio; ITC, indirect treatment comparison; NS, not statistically significant; RR, rate ratio; SS, statistically significant; SUCRA, surface under the cumulative ranking curve.
[a] Hazard Ratio or Rate Ratio as indicated in the treatment/outcome column.
[b] ITT analysis inputs