olaparib versus placebo. In the *BRCAm* subgroup, the HR for TSST/D was 0.44 (95% c.i. 0.29 to 0.67) for olaparib versus placebo; median TSST/D was 23.8 months in the olaparib arm versus 15.2 months in the placebo arm.

AEs occurred more often in the olaparib group, but were largely minor and manageable with dose reductions or interruptions. A greater proportion of patients in the olaparib arm than the placebo arm suffered from severe AEs such as fatigue, anaemia and neutropenia. Serious adverse events (SAEs) occurred in 21.6% of olaparib patients versus 9.7% of placebo patients. These included anaemia, small bowel obstruction, dyspnoea and gastritis. Mortality was slightly higher in the ~~olaparib~~ placebo group than the ~~placebo~~ olaparib group, although the study sample size was too small to conclusively identify any difference in mortality.

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

The ERG has a number of concerns relating to the evidence submitted by the company. In terms of conventional standards of evidence-based medicine, the ERG considers the evidence base to be weak and at a high risk of bias. There are also a number of confounding factors and methodological issues that have not been addressed.

Adaptations to the systematic review were made by the company to bring it in line with the final NICE scope. These were poorly reported in terms of study selection criteria and processes. Clarifications from the company suggest however that the review was well conducted. The ERG believes that all relevant evidence was identified.

Study 19, the only study included in the company's review, has a number of limitations, both methodologically and in terms of its relevance to the decision problem.

Methodological problems include:

- Errors ~~with~~ made when using the interactive voice randomisation system (IVRS) which led to mis-stratification of patients and may account for observed and potentially unobserved imbalances in known and unknown prognostic factors between groups.
- ~~The continuation of some olaparib patients on treatment after disease progression, which contravenes the marketing authorisation for olaparib and would therefore be unlikely to occur in usual practice. This could introduce bias and is likely to confound results for all end points except PFS.~~ Olaparib was not always stopped when RECIST criteria indicated disease progression. It is unclear how "progression of the underlying disease" specified in the marketing authorisation, should be assessed (e.g. radiological, symptomatology, biomarkers),

and therefore unclear if the practice of continuing treatment beyond RECIST progression is in accordance with the licence. However, it would seem that continuation beyond progression by any assessment, especially radiological, could contravene the licence.

- Multiple amendments to the OS analysis plan in terms of interim analyses being added and removed.
- PFS2 was not measured, though this was defined in the final NICE scope and is recommended in EMA guidelines to provide supporting evidence of the persistence of treatment effects in the long-term.
- *Post hoc* addition of a long-term OS analysis leading to the potential introduction of bias.
- *Post hoc* addition of TTD/D, TFST/D and TSST/D leading to the potential introduction of bias.
- Use of the safety analysis dataset rather than the ITT analysis dataset for TDD/D, TFST/D and TSST/D.
- *Post hoc* addition of *BRCA* testing for all patients, including the addition of tumour *BRCA* testing, leading to the potential introduction of bias.
- Crossover of placebo group patients to subsequent PARP inhibitor treatments, leading to the potential introduction of confounding and bias. This issue was addressed in additional analyses provided by the company during the clarification process.
- A lack of clarity regarding when and under what circumstances patients were treated with subsequent chemotherapy after progression.
- The small sample size of the *BRCAm* subgroup, which is reduced further within the crossover analysis which excludes study sites that allowed placebo group crossover.
- Interaction tests were not presented in the CS, and appear to be inconclusive as to whether the *BRCAm* subgroup is statistically significantly different to the rest of the study population.

Problems with relevance to the decision problem and clinical practice in England include:

- The study used both germline (blood test) and tumour (tissue sample test) *BRCA* mutation testing to select patients. These tests are not routinely performed in England. It is unclear whether tumour testing will be possible in England on a large scale. As such, the population who would be treated may differ from that analysed in the subgroup analysis, leading to potential problems with generalisability. It is unclear whether results would have been biased, and if so, whether they would be biased in favour or against olaparib with respect to this issue.
- ~~The study did not use CA-125 to assess progression. This is likely to have lengthened PFS in comparison to practice in England, where CA-125 is used to assess progression in some centres. The extent of the impact of this bias is unknown. It may also mean that patients in~~

~~England would receive treatment for a shorter amount of time on average compared with patients in Study 19. This may impact on both costs and effectiveness in usual clinical practice.~~ Uncertainty around when olaparib will be stopped in clinical practice, and how comparable this will be to practices in Study 19, potentially affecting the generalisability of results.

- The lack of clarity concerning when and under what circumstances patients were treated with subsequent chemotherapy after progression means that TFST/D and TSST/D may not reflect usual clinical practice in England. This may shorten or lengthen the observed TFST/D and TSST/D estimates, though clinical advice received by the ERG suggests that it is more likely that it will shorten estimates, and may affect comparative estimates between study arms.

Table 6: Patient demographics and baseline characteristics for whole population and according to *BRCA* mutation status (reproduced from CS[1] Table 6.5)

	Full analysis set		***BRCAm***	
	Olaparib 400 mg bid (n=136)	**Placebo (n=129)**	**Olaparib 400 mg bid (n=74)**	**Placebo (n=62)**
Age (years)				
Median (range)	58.0 (21–89)	59.0 (33–84)	57.5 (38–89)	55.0 (33–84)
Age group (years), n (%)				
<50	30 (22.1)	20 (15.5)	19 (25.7)	16 (25.8)
≥50 to <65	61 (44.9)	74 (57.4)	38 (51.4)	35 (56.5)
≥65	45 (33.1)	35 (27.1)	17 (23.0)	11 (17.7)
Ethnic population,* n (%)				
Jewish descent				
No	115 (84.6)	112 (86.8)	60 (81.1)	48 (77.4)
Yes	21 (15.4)	17 (13.2)	14 (18.9)	14 (22.6)
ECOG performance status, n (%)				
(0) Normal activity	110 (80.9)	95 (73.6)	62 (83.8)	45 (72.6)
(1) Restricted activity	23 (16.9)	30 (23.3)	11 (14.9)	15 (24.2)
(2) In bed <50% of the time	1 (0.7)	2 (1.6)	0	1 (1.6)
Unknown	2 (1.5)	2 (1.6)	1 (1.4)	1 (1.6)
Primary tumour location				
Ovary	119 (87.5)	109 (84.5)	65 (87.8)	54 (87.1)
Fallopian tube or primary peritoneal	17 (12.5)	19 (14.7)	9 (12.0)	8 (12.9)
Time to progression after completion of penultimate platinum-based regimen				
>6 to ≤12 months	53 (39.0)	54 (41.9)	28 (37.8)	26 (41.9)
>12 months	83 (61.0)	75 (58.1)	46 (62.2)	36 (58.1)
OR to most recent platinum-based regimen				
CR	57 (41.9)	63 (48.8)	36 (48.6)	34 (54.8)
PR	79 (58.1)	66 (51.2)	38 (51.4)	28 (45.2)

*Data are median (range) or number (%); *Ancestry was self-reported*

4.2.1.2 Intervention

The intervention assessed within Study 19 was 400mg olaparib, twice daily (8 *x* 50mg capsules).[24] Patients could be discontinued from treatment at any time at the discretion of the investigators. Retreatment was not permitted. Interruptions and dose reductions were permitted to manage toxicity and AEs. All these are in line with the marketing authorisation for olaparib.[25]

Patients were allowed to continue study treatment following objective progression, provided they were still benefitting and did not meet any other discontinuation criteria (see CS[1] page 65). The licence states that *"It is recommended that treatment be continued until progression of the underlying disease."*[23] However, it is not clear how progression should be assessed, and therefore whether the continuation of treatment contravenes the licence. If it is assumed that radiology is the definitive

measurement of assessment and the licence refers to radiological assessment, the use of olaparib within the study does contravene licencing.

However, it is still important to consider what triggers radiological investigations in clinical practice, as this dictates the time point at which progression will be assessed and detected, and consequently the time point at which treatment will be discontinued in clinical practice, and the generalisability of Study 19. According to the clinical advisors to the ERG, radiological assessment can be triggered by routine radiological monitoring, increases in CA-125, or by symptomatology. There is heterogeneity across the UK in terms of the monitoring offered to patients. For example, in Sheffield, patients will usually be monitored via symptomatology only (symptoms triggers radiological assessment), meaning progression would be detected at a later point in time, and olaparib would be provided for longer, affecting costs and potentially efficacy. Equally, some areas use CA-125 to monitor and trigger radiological assessments, meaning progression would be detected earlier, with inverse potential knock-on effects to costs and efficacy. In Study 19, radiological assessment was conducted 12 weekly up to week 60, and 24 weekly after week 60. It could also be triggered by CA-125 assessment, but does not appear to be triggered by symptomatology. As such, Study 19 does not fully represent clinical practice in England, but the impact of this on estimates of efficacy is unknown. As such, data collected beyond PFS within Study 19 (TFST, TSST, etc.) may be poorly generalisable to clinical practice. This is also likely to have affected the *post hoc* analysis of TTD/D.

~~This is not in line with the marketing authorisation, which states that *"It is recommended that treatment be continued until progression of the underlying disease."*[23] As such, data collected beyond PFS within Study 19 may overestimate relative treatment effects for olaparib compared with clinical practice undertaken in accordance with its licence. This is also likely to have affected the *post hoc* analysis of TTD/D.~~

~~However, there is also an issue around what is considered disease progression in clinical practice. This is not defined in the EPAR23 or the SmPC11 with respect to when treatment with olaparib should be discontinued. Within the trial, PFS was assessed using Response Evaluation Criteria in Solid Tumours (RECIST) criteria. However, progression can also be assessed through CA-125 levels either alone or in conjunction with RECIST (e.g. Study 19 uses this as a secondary outcome). Clinical advisors to the ERG suggested that CA-125 is differentially monitored throughout England, with a substantial minority of patients not receiving this test for monitoring purposes post-chemotherapy. As such, there may be heterogeneity in how "progression" is interpreted in clinical practice. If CA-125 is used as a criterion for assessing progression in clinical practice, this is likely to decrease the duration and potentially the benefit of olaparib treatment. The impact of this potential heterogeneity in clinical practice on the generalisability of the study results is unknown.~~

In addition, it is unclear whether a comparable proportion of patients in Study 19 received subsequent chemotherapy immediately after radiological progression~~, or whether, as is~~as compared to clinical practice~~ usual~~ in England~~,~~, where it is common for patients to only progress~~ed~~ to subsequent chemotherapy once symptoms require~~d~~ treatment. Clinical advice received by the ERG suggests that international practice may be to treat relapse earlier than in the UK, in response to rises in CA-125 and before symptoms present. As such, the event of TFST may be shorter in Study 19 than would be expected in clinical practice in England.

Where toxicity is expected to be increased, as is likely for olaparib given that the alternative is watchful waiting, PFS is required at the least, and survival data should be made available at the time of submission, with long-term follow-up post-approval. However, it is also acknowledged that alternative endpoints may be more appropriate in certain situations.[26]

- OS data are provided up to November 2012, however the latest figures are not provided. The final DCO for OS has not yet been reached [redacted].[9]
- It appears that within Study 19, the EMA guidelines[26] have been adhered to in other respects, and it is likely that the use of TFST/D and TSST/D constitute "alternative endpoints" mentioned above. The CS[1] states that the EMA were consulted in the process of preparing their data for submission. Additionally, clinical advisors to the ERG suggest that TFST/D and TSST/D are clinically relevant outcomes. This is because patients with ovarian cancer in England often only receive subsequent treatment after progression once their symptoms have become problematic. The rationale for this, according to the ERG's clinical advisors, is that the longer the gap between chemotherapies, the better the outcome in terms of patients' quality of life and recovery from toxicity. A rough rule of thumb is that if treatments are given at <6 month intervals, they are less likely to be effective, and that it does not matter at which point in tumour progression any given round of chemotherapy is administered, as it will have the same effect on OS. As such, it could be argued that TFST/D and TSST/D are more clinically relevant than PFS.

However, there are issues with TFST/D and TSST/D specifically in the context of Study 19:

- These analyses were *post hoc*, and were only added once the study results were known.
- Patients were allowed to continue with olaparib after progression at the discretion of the clinician. As noted previously, this ~~contravenes~~ may contravene the marketing authorisation for olaparib. ~~As such, estimates of TFST/D from Study 19 may overestimate relative treatment effects compared with the use of olaparib in accordance with its license.~~ The ERG requested clarification on this issue from the company. Data were provided for the *BRCAm* subgroup only,[9] and are summarised in Table 8. These data indicate that a significant proportion of patients in the olaparib group [redacted]

  This could ~~result in an overestimation~~ affect ~~of treatment effect as measured by PFS,~~ TFST/D, TSST/D and OS as patients may be receiving olaparib for a different amount of time than would occur in clinical practice ~~compared against what would be expected in usual clinical practice~~(see section 4.2.1.2). Data relating to TFST/D, the interval from first subsequent therapy to second subsequent therapy (for the olaparib group only), and survival

(conditional on the point in the pathway rather than according to the point of randomisation) have been used in the company's health economic model (see Chapter 5).

Table 10: Quality assessment results for Study 19 (adapted from CS[1] Table 6.6)

Study question	How is the question addressed in the study?	Grade*	ERG critique – Full population	ERG score*
Was randomisation carried out appropriately?	Eligible patients were randomised in a 1:1 ratio using an IVRS. The investigators/ sites determined the appropriate stratification variables for each patient at the time of randomisation. A blocked randomisation was generated and all centres used the same list in order to minimise imbalance in numbers of patients assigned to each group.	Yes	The randomisation plan was appropriate, as stated in the CS[1] However, it is evident from the CSR that the randomisation process was not performed as planned, due to mis-stratification by the IVRS.	No (High risk)
Was the concealment of treatment allocation adequate?	The active and placebo capsules were identical and presented in the same packaging to ensure blinding of the study medication.	Yes	Inappropriately scored and supported. Allocation concealment is only partly assured by identical treatment presentation. Data from Ledermann *et al*, 2014[22] confirms that concealment of treatment allocation was adequate.	Yes (Low risk)
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The demographic characteristics of the *BRCAm* patients were generally consistent with the overall population and the two treatment groups were well balanced in terms of age, race and ethnicity. The age distribution of the *BRCA* population was younger than the overall population but this is consistent with the hereditary nature of *BRCAm*. At diagnosis, the majority of patients had a tumour that was FIGO Stage IIIC (59%). There was a slight imbalance between groups in a number of patients with ECOG PS 0 or 1. For response to previous platinum therapy, a lower percentage in the olaparib compared with placebo group had a CR, and vice versa for PR.	Yes	Inappropriately scored and supported. Imbalances in treatment arms are described and may have confounded results. Some analyses included adjustments for the three stratification factors, but not for the imbalance in~~other imbalances~~ ECOG performance status. It was not clear if all *BRCAm* analyses adjusted for stratification factors.	No (high risk)

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Blinding was maintained throughout the study unless in the case of medical emergency, where un-blinding was necessary. The active and placebo capsules were identical and presented in the same packaging. Un-blinding did not occur until after all planned analyses.	Yes	21.1% of patients in total were unblinded. In some cases, this may have affected outcome measurement.	No (High risk)
Were there any unexpected imbalances in dropouts between groups? If so, were they explained or adjusted for?	There were no unexpected imbalances in dropouts between treatment groups.	No	Appropriately scored and justified.	No (Low risk)
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes measures are accounted for.	No	The statement that *"all outcome measures are accounted for"* does not directly answer the question of whether they were reported. In the context of the NICE scope,[10] all outcomes that were measured were reported.	No (Low risk)
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Efficacy data from this study was summarised and analysed on an ITT basis using randomised treatment. Analyses were undertaken for the overall study population and for patients with *BRCA*-mutated ovarian cancer.	Yes	An ITT analysis was performed according to patient randomisation. Patients who withdrew consent were censored and early censoring was balanced between groups.	Yes (Low risk)

**Grade scored as yes/no/not clear or N/A*

There are also a number of protocol changes which may have affected the estimates of efficacy produced from the study. These include:

- Multiple amendments to the timing of OS analyses
- The addition of *post hoc* outcomes TFST/D, TSST/D and TTD/D
- *Post hoc* changes to how *BRCA* status was determined. At study outset, no *BRCA* testing was planned. In a *post hoc* amendment, after some evidence of enhanced efficacy in *BRCA* patients was observed, testing was implemented for all patients, including not only germline but also tumour *BRCA* status.
- The *post hoc* addition of all study outcomes to the *BRCA* subgroup analysis. It is unclear which outcomes were added to the plan in the round of changes that introduced the *BRCA* subgroup in June 2010.

Patient spectrum

The ERG notes the following concerns regarding the patient spectrum within Study 19:

- The population within Study 19 was wider than the population specified within the marketing authorisation for olaparib;[11] all supporting data are subgroup data from Study 19.
- A greater proportion of patients in the olaparib arm had ECOG performance status "normal" than in the placebo arm. Results were not adjusted for this factor.
- The *BRCAm* subgroup was generally younger than the whole population, which may bias results. The direction of this potential bias is unknown.
- A greater proportion of patients in the olaparib arm had a PR to their most recent platinum-based chemotherapy compared against placebo patients; this may bias results which have not been adjusted for this factor (some outcomes were adjusted, but it is not always clear which) in an unknown direction, but possibly in favour of olaparib based on subgroup analyses presented in Figure 6.6 of the CS.

Intervention

- Patients were allowed to continue on study treatment following objective progression, which ~~is not in line with the wording of the marketing authorisation for olaparib~~may or may not reflect how olaparib will be used in clinical practice.[11] This may have confounded any outcome recorded beyond PFS.
- It is unclear whether the definition of disease progression that dictated treatment termination and criteria that dictated subsequent treatment were, on average, comparable to clinical practice in England. These potential biases could operate in either direction, depending on whether CA-125 was used more or less frequently than in clinical practice.

Outcomes

- PFS is a proxy for OS. Whilst its use is acceptable for regulatory purposes, as it demonstrates some benefit to patients, it is not a substitute for OS.
- The study was not powered to detect a statistically significant difference in OS. This means that an absence of an OS advantage could reflect a type II error.
- TTD/D, TFST/D and TSST/D were all *post hoc* analyses.
- TTD/D is not a measure of real-world efficacy as patients were allowed to continue treatment after progression, and is not a measure of real-world drug use for the same reason.
- Crossover of placebo patients to PARP inhibitor treatment may have resulted in the confounding of outcomes recorded after PFS, as placebo patients receive the benefit of the PARP inhibitor. Crossover may also have caused patients in the placebo arm who enrolled in subsequent PARP inhibitor trials to have received either their first or second subsequent therapy earlier than would be expected in usual clinical practice, depending on eligibility and treatment criteria of the trial.

Table 18 summarises the factors which may have affected the estimates of efficacy produced in Study 19.

Table 18: Summary of factors that may affect the estimates of efficacy produced in Study 19

Item	Issue	Likely effect in opinion of the ERG	Likely direction in opinion of the ERG
Systematic review	Systematic review irregularities	• None	None
Study design	Use of subgroup data only to support submission	• Possibility of a Type I error	May favour olaparib
	Subgroup defined after the study commenced	• Risk of data mining having occurred	May favour olaparib
	Interaction tests inconclusive	• It is not clear whether the *BRCAm* subgroup is different to the whole population	Unclear
	Lack of consistent definition of "*BRCAm*" patients – initially only those with known status, study design changed to include testing of all patients and inclusion of tumour *BRCAm* patients as well as germline patients.	• Increase sample size to gain statistically significant results • Constitutes *post hoc* analysis	Increase statistical power
	Phase II clinical trial	• Results are unconfirmed	May favour olaparib
	Unclear if ~~subsequent treatment immediately followed progression, which does not always occur in England~~ timing of chemotherapy after progression reflects clinical practice	• Unclear, could shorten estimates of TFST/D and TSST/D compared with clinical practice	May disadvantage olaparib
	Randomisation IVRS mis-stratified patients	• Imbalances in groups with unknown impact on efficacy	Unclear
	Un-blinding occurred in around 20% patients after progression	• Hastened TFST/D or TSST/D • Affected other outcomes post-PFS	Unclear
	Multiple changes to study protocol and analysis plan	• Risk of data mining having occurred • Risk of manipulation of outcome times to favour drug	May favour olaparib
	Protocol deviations	• Various	Unclear
Patients	ECOG performance imbalance	• Bias in unknown direction	Unclear
	BRCA subgroup younger than whole population	• May be a confounder in *BRCA* subgroup analysis	Unclear
	PR/CR imbalance between olaparib and placebo arms in whole group and *BRCA* group	• Bias in unknown direction for outcomes where no adjustment for PR/CR	Unclear, may favour olaparib for PFS in *BRCAm* subgroup
	PR/CR imbalance between *BRCA* group and whole population	• May be a confounder in *BRCAm* subgroup analysis	Unclear

Item	Issue	Likely effect in opinion of the ERG	Likely direction in opinion of the ERG
Intervention	Continuation of study treatment post-progression.	• Confounding of outcomes beyond PFS~~, likely to cause over-estimation of efficacy~~ • Especially likely to affect TTD/D, which ~~will~~ may not reflect real-world clinical effectiveness or drug use	May favour olaparib
	Unclear if "progression" defined in similar way to clinical practice in England, which impacts how long treatment given for	• Unclear, could affect generalisability and efficacy	Unclear
	Unclear if subsequent treatment administered according to same criteria as in clinical practice in England	• Unclear, could affect generalisability and efficacy	Unclear
Outcomes	PFS is a proxy for OS	• Unclear if there is an OS benefit	Favours olaparib at PFS assessment point
	Study not powered to detect difference in OS	• Cannot rule out a type II error in the absence of OS advantage being demonstrated	Favours olaparib at PFS assessment point
	TFST/D, TSST/D and TTD/D outcomes defined *post hoc*	• High risk of bias/ data mining having occurred • Not all outcomes requested by NICE	May favour olaparib
Crossover	Crossover of placebo patients to PARP inhibitor therapy post-progression	• May truncate estimates of TFST/D and TSST/D in placebo arm, favouring olaparib	May favour olaparib
		• May overestimate OS and TSST/D in the placebo arm, disadvantaging olaparib	Disadvantages olaparib

in the olaparib arm versus 4.6 months in the placebo arm. It is unclear what this outcome actually demonstrates given the confounding issues.

TFST/D was considered by the ERG to be a *post hoc* analysis potentially confounded by patient switching and un-blinding of some study participants. Within the whole population, the HR for TFST/D was 0.41 (95% c.i. 0.31 to 0.54). The HR for TFST/D was 0.33 (0.22 to 0.50) for olaparib versus placebo within the *BRCAm* subgroup; median TFST/D within the *BRCAm* subgroup was 15.6 months in the olaparib arm versus 6.2 months in the placebo arm. This outcome suggests that patients receiving olaparib commenced their first subsequent therapy later than patients on placebo, but it should be borne in mind that the outcome was a *post hoc* analysis and may be confounded in either direction.

TSST/D was considered by the ERG to be a *post hoc* analysis potentially confounded by patient switching and un-blinding of some study participants. In the whole population, the HR for TSST/D was 0.54 (95% c.i. 0.41 to 0.72) for olaparib versus placebo. In the *BRCAm* subgroup, the HR for TSST/D was 0.44 (95% c.i. 0.29 to 0.67) for olaparib versus placebo; median TSST/D was 23.8 months in the olaparib arm versus 15.2 months in the placebo arm. This outcome suggests that patients receiving olaparib commenced their second subsequent therapy later than patients on placebo, but it should be borne in mind that the outcome was a *post hoc* analysis and may be confounded in either direction.

Quality of life, as measured by FACT-O, and the indexes TOI and FOSI, was not negatively impacted by therapy with olaparib compared with therapy with placebo. Putative HRQoL gains are likely to be a consequence of avoided rounds of chemotherapy, however there is no direct evidence to support this.

AEs occurred more often in the olaparib group, but were largely minor and manageable with dose reductions or interruptions. A greater proportion of patients in the olaparib arm than the placebo arm suffered from severe AEs such as fatigue, anaemia and neutropenia. SAEs occurred in 21.6% of olaparib patients versus 9.7% of placebo patients. These included anaemia, small bowel obstruction, dyspnoea and gastritis. Mortality was slightly higher in the placebo group than the olaparib group, although the study sample size was too small to conclusively identify any difference in mortality. The ERG notes that:

- Although olaparib does not seem to impact on quality of life, and may confer some OS benefit, it does have a worse AE profile than placebo, including a higher occurrence of serious and severe AEs.

Table 46: Central estimates of cost-effectiveness presented within the CS

Central estimates of cost-effectiveness (point estimates of parameters)							
Option	LYGs	QALYs	Costs	LYGs	Inc. QALYs	Inc. costs	ICER
Olaparib	3.55	2.58	£82,041	1.17	0.89	72,143	**81,063**
Routine surveillance	2.38	1.69	£9,898	-	-	-	-
Central estimates of cost-effectiveness (expectation of the mean)							
Option	LYGs*	QALYs	Costs	LYGs*	Inc. QALYs	Inc. costs	ICER
Olaparib	-	2.60	£82,048	-	0.90	£72,232	**£79,953**
Routine surveillance	-	1.70	£9,816	-	-	-	

Inc. – incremental; LYG – life year gained; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio
**life years gained are not recorded within the company's PSA sub-routine*

Based on the probabilistic analysis of the model, olaparib is expected to produce an additional 0.90 QALYs at an additional cost of £72,232 compared against routine surveillance. The ICER for olaparib versus routine surveillance is therefore expected to be £79,953 per QALY gained. The results based on point estimates of parameters are similar, with olaparib yielding an ICER of £81,063 per QALY gained compared against routine surveillance.

Table 47 presents a breakdown of costs and QALYs for olaparib and routine surveillance based on values reported within the company's model (deterministic, discounted and half-cycle corrected).

Table 47: Breakdown of costs and QALYs for olaparib and routine surveillance

Component	**Olaparib**	**Routine surveillance**	**Difference (olaparib vs routine surveillance)**
QALYs			
Progression-free on treatment	1.24	0.54	0.70
Progression-free discontinued treatment	0.52	0.21	0.31
First subsequent therapy	0.39	0.44	-0.05
Second subsequent therapy	0.44	0.50	-0.06
Total QALYs gained	2.58	~~2~~1.69	0.89
Costs			
Maintenance therapy (olaparib)	£70,152.48	£0.00	£70,152.48
AEs	£168.20	£59.23	£108.98
First subsequent therapy	£1,459.07	£1,887.95	-£428.88
Second subsequent therapy	£1,423.16	£1,669.23	-£246.07
Follow-up	£5,498.29	£2,761.14	£2,737.15
End-of-life	£3,339.40	£3,520.07	-£180.67
Total costs	£82,040.60	£9,897.60	£72,142.99

QALY – quality-adjusted life year

The results presented in Table 47 suggest that the majority of the incremental QALY gains for olaparib versus routine surveillance are generated within the "progression-free" states (1.01 additional

Element of HTA	Reference Case	ERG comments
Defining the decision problem	The scope developed by NICE	The company's economic analysis is generally in line with the final NICE scope. It is noteworthy that the outcomes data used in the model do not include OS or PFS.
Comparator(s)	As listed in the scope developed by NICE	The company's model includes routine surveillance ("watch and wait") as the comparator.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	The base case analysis includes direct health effects on patients and costs borne by the NHS and PSS. The wider impact of expanding *BRCA* mutation testing to unaffected relatives is included in the company's secondary analysis. The model includes functionality to consider societal costs although these are not presented within the CS.
Perspective on costs	NHS and PSS	
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The model estimates the incremental cost per QALY gained for olaparib versus routine surveillance.
Time horizon	Long enough to reflect all important differences between the technologies being compared	A 15-year time horizon is assumed, which is intended to reflect the patients' remaining lifetime. Approximately 1.46% patients in the olaparib group are still alive at this timepoint. The model does not include the functionality to consider longer time horizons.
Synthesis of evidence on health effects	Based on systematic review	All outcomes data used in the model are based on analyses of data from the *BRCAm* subgroup within Study 19.[27] This was the only study identified within the company's systematic review of clinical effectiveness evidence.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults	Health outcomes are measured and valued in terms of QALYs. Health utility scores for the progression-free states are based on a mapping of FACT-O data collected in Study 19 to the EQ-5D based on an algorithm reported by Longworth *et al*.[46] Health utilities for the first- and second-subsequent chemotherapy states are taken from OVA-301 trial.[38,48] All utilities were valued using the UK tariff.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No equity weightings are applied.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Costs are valued from an NHS and PSS perspective.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Health outcomes and costs are discounted at a rate of 3.5% per annum.

The company's health economic analysis is generally in line with the NICE Reference Case[44] and the final NICE scope.[10] As discussed in Section 5.2, the company's model adopts a 15-year time horizon. However, 1.46% patients in the olaparib group are still alive at this timepoint. The model does not include functionality to consider longer time horizons. It is very unlikely however that the adoption of a longer time horizon would have a marked impact upon the company's ICER for olaparib versus routine surveillance. It is noteworthy that whilst specified in the final NICE scope,[10] the company's base case analysis does not include the costs of *BRCA* mutation testing; this is discussed further in Section 5.3.3.

5.3.3 Summary of main issues identified through critical appraisal of the company's model

The main issues identified through the critical appraisal of the company's health economic analysis and the model upon which this is based are summarised in Box 2.

Box 2: Summary of main issues identified through the critical appraisal of the company's model

(1) Model errors and other issues surrounding model implementation
(2) Potential bias in the *BRCAm* subgroup within Study 19
(3) Concerns regarding company's model structure and use of outcomes data from Study 19
(4) Potential confounding of end points used in the company's model
(5) Concerns regarding the methods and presentation of modelling of time-to-event outcomes
(6) Discordance between model predictions and observed data
(7) Issues surrounding HRQoL within the company's model
(8) Omission of the cost of *BRCA* mutation testing from the company's base case analysis
(9) Issues surrounding the company's secondary analysis of the cost-effectiveness of olaparib including the wider costs and benefits of *BRCA* mutation testing for unaffected relatives
(10) Limited use of sensitivity analysis

(1) Model errors and other issues surrounding model implementation

The ERG partially rebuilt the company's model in order to assess the logic of the company's model structure, to draw out any unwritten assumptions and to identify any errors in the implementation of the model. The semi-Markov trace for the olaparib and routine surveillance groups was fully redeveloped using time-to-event data contained within the "Survival data" worksheet of the company's model. A comparison of intermediate model outputs from the company's model and the ERG rebuilt model is presented in Table 51. The ERG rebuilt model produced estimates of 2.58 discounted QALYs for olaparib and 1.69 discounted QALYs for routine surveillance; these are the same as those generated from the deterministic version of the company's model. The ERG is satisfied

and 0.65 relate to states of "progression-free survival" and "progressed disease" rather than states of "first subsequent therapy" and "second subsequent therapy", respectively. In reality, patients receiving chemotherapy would have a progression-free period and a post-progression period and it is likely that each of these states would be associated with different levels of HRQoL. The ERG considers that the estimates used do not fully reflect the health states included in the company's model but also recognise the lack of alternative relevant preference-based estimates within the literature.

(8) Omission of the cost of BRCA mutation testing from the company's base case analysis

The final NICE scope[10] states that *"The economic modelling should include the cost associated with the diagnostic testing for BRCA1/2m in people with OC who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test."* In addition, the NICE Guide to the Methods of Technology Appraisal[44] states that *"If a diagnostic test to establish the presence or absence of this biomarker is carried out solely to support the treatment decision for the specific technology, the associated costs of the diagnostic test should be incorporated into the assessments of clinical and cost effectiveness. A sensitivity analysis should be provided without the cost of the diagnostic test."* Whilst the CS[1] includes a sensitivity analysis in which the cost of *BRCA* mutation testing is included, the company's base case analysis excludes the cost of the diagnostic test. The inclusion of the cost of *BRCA* mutation testing increases the company's ~~probabilistic~~ deterministic base case ICER for olaparib versus routine surveillance from £81,063 per QALY gained to £84,326 per QALY gained.

(9) Issues surrounding the company's secondary analysis of the cost-effectiveness of olaparib including the wider costs and benefits of BRCA mutation testing for unaffected relatives

The final NICE scope[10] does not indicate that the potential additional health benefits of *BRCA* mutation testing should be included in the economic analysis. In the event that these costs and benefits are deemed relevant to decision-making, the ERG has both practical and theoretical concerns regarding their inclusion.

The ERG has several concerns regarding the appropriateness of combining the results of the company's model and the model developed to inform NICE CG164.[8] Firstly, there are differences between the two models in terms of the treatment pathways assumed for ovarian cancer; specifically, the guideline model does not include olaparib as a treatment option, hence the company's analysis reflects a situation in which *BRCA*-testing and olaparib treatment are available for the index case, but the treatment is not available for relatives; this is somewhat inconsistent. Secondly, the ERG notes that the use of five family pedigrees, combined with average costs and QALYs for unaffected family members, is limited and may not adequately reflect the range of possible family structures within the population under consideration. It is therefore unlikely that the average ICER across the five

Methods used to address question 1: What is the expected incremental survival gain for olaparib versus placebo?

Based on the extrapolation of OS detailed in the previous section, the mean survival in each group was estimated using a simple restricted means approach by estimating the AUC for each treatment group over a time period of 15-years (the company's modelled time horizon). Mean incremental survival gains were estimated as the difference in the 15-year restricted mean AUC between the two groups. Mean survival was estimated using the generalised F, generalised gamma, gamma, log normal, log logistic and Weibull models fitted to the empirical OS data. Restricted means estimates were not discounted.

Methods used to address question 2: What is the expected incremental QALY gain for olaparib versus placebo?

A simple partitioned survival model was developed incorporating four health states: (1) alive, not yet progressed to first subsequent therapy, on maintenance therapy; (2) alive, not yet progressed to first subsequent therapy, discontinued maintenance therapy (3) alive, following progression to first subsequent therapy, and (4) dead.

Estimates of the cumulative probability of survival up to time *t* were estimated directly from each potentially ~~plausible~~ suitable candidate parametric survivor function fitted within the ERG's survival modelling exercise detailed above.

The probabilities of not experiencing an event at each timepoint were adjusted to account for two logical inconsistencies which may have resulted from the extrapolation process:

(i) The outcome of initiation of first subsequent therapy or death for a given patient must always occur before, or at the same time as, death due to any cause. If the probability of initiation of first subsequent therapy at time *t* was greater than the probability of OS at time *t*, the probability of initiation of first subsequent therapy or death at time *t* was set to zero.

(ii) The outcome of treatment discontinuation or death for a given patient must always occur before, or at the same time as, the initiation of first subsequent therapy or death. If the probability of treatment discontinuation or death at time *t* was greater than the probability of initiation of first subsequent therapy or death at time *t*, the probability of treatment discontinuation or death at time *t* was set to zero. A similar adjustment to TTD/D was made within the company's model.

The probability of residing in each of the model health states at each timepoint was calculated as follows.

The probability of being alive, not having progressed to first subsequent therapy, and being on maintenance therapy (i.e. in State 1) at time *t* was calculated as:

$S(t_i)_{TTD}$ [i]

The probability of being alive, not having progressed to first subsequent therapy, having discontinued maintenance therapy (i.e. in State 2) at time *t* was calculated as:

$S(t_i)_{TFST} - S(t_i)_{TTD}$ [ii]

The probability of being alive, following progression to first subsequent therapy (i.e. in State 3) at time *t* was calculated as:

$S(t_i)_{OS} - S(t_i)_{TFST}$ [iii]

The probability of being dead (i.e. in State 4) at time *t* was calculated as:

$1-S(t_i)_{OS}$ [iv]

A half cycle correction was applied to adjust for the timing of events. Mean QALYs in each treatment group were estimated by applying utility values drawn from the company's base case analysis.[1] As a single utility value was not available for patients in State 3 (progression to chemotherapy), a simple mean of the two utility values reported by the manufacturer in TA222[38] was assumed (utility=0.68). Health outcomes were discounted at a rate of 3.5% per annum.[44]

The model was analysed across 108 scenarios which reflect the combinations of the most ~~plausible~~ suitable curves fitted within the ERG's survival modelling exercise (see Table 63). In each scenario, the same survivor function was applied to both treatment groups.