Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Bevacizumab in combination with paclitaxel and carboplatin for the first-line treatment of ovarian cancer

Produced by Southampton Health Technology Assessments Centre

Authors Keith Cooper, Senior Research Fellow, SHTAC
Karen Pickett, Research Fellow, SHTAC
Geoff Frampton, Research Fellow, SHTAC
Vicky Copley, Research Fellow, SHTAC
Jackie Bryant, Principal Research Fellow, SHTAC

Correspondence to Jackie Bryant
Southampton Health Technology Assessments Centre
University of Southampton
First Floor, Epsilon House
Enterprise Road, Southampton Science Park
Southampton SO16 7NS

Date completed 23 October 2012

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 10/84/01.

Declared competing interests of the authors

None.

Acknowledgements

We are very grateful to the following expert who offered clinical advice and comments on the draft report: Dr Clare Green, Consultant Oncologist, Southampton University Hospitals NHS Trust, Southampton General Hospital, Tremona Road, Southampton, Hampshire, SO16 6YD.

We also thank: Karen Welch, Information Specialist, SHTAC, for commenting on the manufacturer's search strategy, and Jeremy Jones, Principal Research Fellow, SHTAC, for acting as internal editor for the ERG report.

Rider on responsibility for report

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

This report should be referenced as follows:

Cooper K, Pickett K, Frampton GK, Copley V, Bryant J. Bevacizumab in combination with carboplatin and paclitaxel for the first-line treatment of ovarian cancer. A Single Technology Appraisal. SHTAC. 2012

Contributions of authors

K Cooper (Senior Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; K Pickett (Research Fellow) critically appraised the clinical effectiveness systematic review and drafted the report; G K Frampton (Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; V Copley (Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; J Bryant (Principal Research Fellow) critically appraised the clinical effectiveness systematic review, drafted the report and project managed the review.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

TABLE OF CONTENTS

LIST OF TABLES

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

LIST OF ABBREVIATIONS

ADS	Abdominal discomfort score
AE	Adverse event(s)
CA-125	Serum cancer antigen 125
CEA	Cost effectiveness analysis
CHMP	Committee for Medicinal Products for Human Use of the European Medicines Agency
CPB15	Carboplatin, paclitaxel, bevacizumab initiation
CPB15+	Carboplatin, paclitaxel, bevacizumab maintenance dose 15mg/kg
CPP	Carboplatin, paclitaxel, placebo
CPB7.5+	Carboplatin, paclitaxel, bevacizumab maintenance dose 7.5mg/kg
CRD	Centre for Reviews and Dissemination
DSA	Deterministic Sensitivity Analysis
EMA	European Medicines Agency
ERG	Evidence Review Group
FACT-O TOI	Functional Assessment of Cancer Therapy-Ovary Trial Outcome Index
GOG	Gynaecologic Oncology Group
FIGO	International Federation of Gynaecologists and Obstetricians
HRQoL	Health Related Quality of Life
ICER	Incremental Cost Effectiveness Ratio
IRC	Independent Review Committee
ITT	Intention to Treat
LY	Life years
LYG	Life year gained
MS	Manufacturer's Submission
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NPT	Non-protocol therapy (prior to disease progression)
OR	Odds Ratio
ORR	Objective response rate
OS	Overall survival
PD	Disease progression
PFS	Progression free survival
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
q3w	Every 3 weeks
QALY	Quality Adjusted Life Year
QoL	Quality of Life
RCT	Randomised Controlled Trial
RR	Relative Risk
SLD	Sum of Longest Diameter
SPC	Summary of Product Characteristics
WTP	Willingness to pay

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

SUMMARY

Scope of the manufacturer submission

The manufacturer's submission (MS) reflects the scope of the appraisal issued by the National Institute for Health and Clinical Excellence (NICE). This was to consider bevacizumab in combination with paclitaxel and carboplatin for the first-line treatment of ovarian cancer.

Summary of submitted clinical effectiveness evidence

The clinical effectiveness evidence in the MS comes from two RCTs of bevacizumab in combination with carboplatin and paclitaxel for the first-line treatment of advanced ovarian cancer:

- A key RCT, GOG-0218, is a double-blind placebo controlled trial of bevacizumab at a dose of 15mg/kg q3w for up to 15 months in a population of Stage III and Stage IV patients;
- A supporting trial, ICON7, is an open label study of bevacizumab at a dose of 7.5mg/kg q3w for up to 12 months in patients with Stage I – IV disease.

The primary outcome is progression free survival (PFS). From the GOG-0218 study there was a statistically significant improvement in the median PFS of 6 months in the trial arm comprising bevacizumab in combination with chemotherapy followed by bevacizumab maintenance therapy (CPB15+) arm compared to the chemotherapy plus placebo (CPP) arm (CPP 12 months, CPB15+ 18 months; HR 0.645, 95%CI 0.551, 0.756, $p<0.001$). No difference in PFS was found for patients who received bevacizumab in combination with chemotherapy without subsequent bevacizumab maintenance therapy (CPB15) compared with patients who received chemotherapy alone. The observed PFS benefit in the CPB15+ group was also shown across subgroups by disease stage and debulking status.

Secondary outcomes include overall survival (OS), objective response rate (ORR) and health related quality of life (HRQoL). Results from the GOG-0218 study suggest that there are no statistically significant differences between the treatment arms for OS and HRQoL, although the independent review committee assessment found statistically significant differences for ORR. Adverse events (AEs) for which the incidence was ≥ 10% higher in the bevacizumab-containing

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

arms than the chemotherapy-alone arm were stomatitis, dysarthia, headache, epistaxis, and hypertension, and were consistent with the known safety profile of bevacizumab.

Results from the ICON7 trial are only presented in an appendix here because this study does not match the patient population specified in the scope (except through subgroup analysis of patients with advanced disease). Also, although the reduced dose of bevacizumab used in the trial more closely reflects UK clinical practice, it is not licensed.

Summary of submitted cost effectiveness evidence

The manufacturer's submission to NICE includes:

i) a review of published economic evaluations of bevacizumab in advanced or metastatic ovarian cancer.

ii) a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of bevacizumab in combination with carboplatin and paclitaxel compared with carboplatin and paclitaxel for first-line line treatment in women with newly diagnosed stage III or IV ovarian cancer.

The manufacturer conducted a systematic search of the literature to identify economic evaluations of bevacizumab in advanced or metatstatic ovarian cancer from a UK perspective. Three studies were included for full review and the cost effectiveness results presented.

The cost effectiveness analysis (CEA) uses a 3 state semi-Markov model to estimate the cost-effectiveness of bevacizumab in combination with carboplatin and paclitaxel (CPB) compared with carboplatin and paclitaxel (CP) for first-line treatment in women with newly diagnosed stage III or IV ovarian cancer. The model has health states for PFS, progression and death. The model adopted a time horizon of ten years with a cycle length of 1 week. The model costs and outcomes were discounted at 3.5% per annum. The perspective of the model is the UK NHS and Personal Social Services (PSS) and results are presented as incremental cost per QALY gained.

The transitions of patients from PFS to disease progression and death were derived from the GOG-0218 and ICON7 RCTs. HRQoL is used in the model for the health states for PFS and progression, based upon EQ-5D surveys of patients in the ICON7 study. Resource use in each

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

health state was based on a previous NICE appraisal in ovarian cancer and costs were taken from the BNF63, PSSRU and NHS reference costs 2010/11.

The model was validated by comparing OS predicted by the model with an external publication using ovarian cancer patients with similar disease severity and surgical outcomes.

The results from the economic evaluation are presented based on the GOG-0218 (and ICON7) RCTs. Results are presented as incremental cost per QALY gained for CPB vs. CP. For the base case an incremental cost per QALY gained of £144,066 is reported for GOG-0218 (and £31,592 for ICON7).

Sensitivity analyses were presented for a limited number of parameters. The key drivers of the cost effectiveness results are the dose and cost of bevacizumab and the duration of the treatment. The manufacturer conducted a probabilistic sensitivity analysis (PSA) which showed there is a 0% probability of CPB being cost-effective, relative to treatment with CP, at a threshold willingness to pay of £30,000 per QALY gained, based upon the GOG-0218 RCT, (and 42% based upon the ICON7 RCT).

Commentary on the robustness of submitted evidence

Strengths

- The MS contains systematic searches for the clinical and cost effectiveness studies of bevacizumab. It appears unlikely that these have missed any studies that would have met the inclusion criteria.
- The systematic review meets the Centre for Reviews and Dissemination (CRD) criteria for methodological quality.
- The economic model presented in the MS used an appropriate approach for the disease area.
- The model has used an appropriate methodology to calculate PFS, based upon the relevant RCT for this treatment.
- All relevant costs and resources have been included and are transparently calculated.
- The cost effectiveness analysis meets the requirements of the NICE reference case.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Weaknesses and Areas of uncertainty

- The MS does not report details of the process used to conduct the systematic review although meeting criteria for methodological quality.
- Results from the key RCT which meets the NICE scope are presented for different assessments and different time points (with and without censoring for method of disease progression measurement) producing a range of hazard ratios. Although the results are similar it is not clear what the actual size of effect is likely to be.
- The licensed dose used in the key trial is not the same as that used in current clinical practice.
- In clinical practice bevacizumab is given to patients with Stage III residual disease only which is a subset of patients within the key trial.
- The evidence also includes an RCT which more closely reflects current clinical practice in terms of dosage of bevacizumab; however, this dosage is not licensed and the patient population in the trial does not match the NICE scope except through subgroup analysis.
- It should be noted that the MS does not report p-values in some analyses, making interpretation of findings difficult as it is not possible to tell the level at which differences between groups are statistically significant. (Clarification requested from the manufacturer stated that outstanding p values were not available.)
- The MS has not included all model parameters in either the deterministic or probabilistic sensitivity analyses and so the full uncertainty around the model results has not been shown. In particular key parameters associated with clinical effectiveness (PFS) and treatment costs have been omitted.
- The treatment duration used within the model has been underestimated by using a maximum of one year, rather than 15 months as stated in the Summary of Product Characteristics (SPC) for bevacizumab and for the GOG-0218 trial, and therefore the cost of bevacizumab has been underestimated.
- The model presents results using a shorter than expected time horizon of ten years.

Summary of additional work undertaken by the ERG

The ERG conducted additional analyses for changes to the bevacizumab treatment duration, time horizon, and the cost of bevacizumab (assuming lower dosage).

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Results show that using a time horizon of 25 years with treatment for a maximum of 15 months the cost effectiveness of CPB compared with CP was £142,477 (i.e. similar to the base case results). Using a reduced treatment cost, equivalent to a lower dosage of 7.5 mg/kg, as often used in clinical practice, and assuming no change in the treatment efficacy, the incremental cost effectiveness ratio (ICER) of CPB compared with CP reduced to £77,884.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

1 Introduction to ERG Report

This report is a critique of the manufacturer's submission (MS) to NICE from Roche on the clinical effectiveness and cost effectiveness of bevacizumab in combination with paclitaxel and carboplatin for the first-line treatment of ovarian cancer. It identifies the strengths and weaknesses of the MS. A clinical expert was consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the MS was requested from the manufacturer by the ERG via NICE on 6 September 2012 (sent to the manufacturer on 19 September 2012). A response from the manufacturer via NICE was received by the ERG on 3 October 2012 and this can be seen in the NICE evaluation report for this appraisal.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The MS provides a clear and accurate overview of ovarian cancer.

2.2 Critique of manufacturer's overview of current service provision

The MS provides an accurate overview of current service provision.

2.3 Critique of manufacturer's definition of decision problem

Population

The population described in the decision problem is appropriate for the NHS.

Intervention

The description of the intervention in the decision problem is appropriate for the NHS. The product was granted marketing authorisation in December 2011. The licensed dose of bevacizumab, in combination with carboplatin and paclitaxel, is 15mg/kg body weight given every 3 weeks by intravenous infusion. The dose used in clinical practice in the UK is 7.5mg/kg as first-line treatment for Stage III sub-optimally debulked patients.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Comparators

The main comparator in the MS decision problem is combination chemotherapy using carboplatin with paclitaxel, which represents one of the recommended platinum-based treatment options for advanced ovarian cancer in the UK.

Outcomes

The outcomes included in the MS are appropriate and clinically meaningful to patients.

Economic analysis

The economic evaluation in the MS decision problem appears to be appropriate, being a cost utility analysis from the NHS and Personal Social Services (PSS) perspective.

Other relevant factors

Subgroups reported in the MS include analysis by disease stage and debulking status and by various baseline risk factors.
(Clarification requested from the manufacturer on whether subgroups were prespecified and powered statistically stated that no power calculations are available for pre-specified subgroups)

The MS states that issues relating to equity or equality are not applicable and this is in line with the decision problem in the NICE scope.

3 CLINICAL EFFECTIVENESS

3.1 Critique of manufacturer's approach to systematic review

3.1.1 Description of manufacturer's search strategy

The manufacturer's literature searches were checked by an information specialist. Overall, the searches are adequately documented by the manufacturer, with satisfactory database selection. The search strategy comprises the use of free text and index terms, appropriately combined. There were a few minor inconsistencies and errors, however not deemed significant enough to miss vital evidence. Differing host systems and syntax employed between the ERG and manufacturer dictates that the searches could not be exactly replicated. A suitable selection of

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

conferences were cited as searched by the manufacturer both electronically and hand-searched. The ERG information specialist undertook some additional searches as follows: a fuller RCT filter search was applied to Medline and Embase. On-going trials searches were undertaken on the following clinical trials registries: UKRCN Study Portfolio, controlled-trials.com, clinical trials.gov and WHO ICTRP. The results were checked by an ERG researcher. No additional trials identified were relevant to the decision problem.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection.

The MS clearly states the inclusion and exclusion criteria in Appendices 1 (MS p.214) and 6 (MS p.231) of the submission. The inclusion criteria reflect the final scope issued by NICE and the licensed indication; that is to include studies of patients with advanced ovarian cancer. However, the manufacturer did not specify dose of bevacizumab as an inclusion or exclusion criterion. This means that studies did not have to use the licensed dosage of bevacizumab (15mg/kg) to be included in the review. Therefore, the inclusion criteria are wider than the scope and the licensed indication.

Study quality and setting were not stated as inclusion or exclusion criteria, and this reflects the final scope. Separate sets of criteria were used to assess RCT (MS p.214) and non-RCT studies (MS p.231) identified from the searches for inclusion. In the RCT inclusion and exclusion criteria, the only study design limitations were that phase I studies, non-RCT studies and reviews were excluded. In the non-RCT inclusion and exclusion criteria, studies that included fewer than 200 patients and RCT studies were excluded. The manufacturer does not provide a justification for excluding non-RCT studies with fewer than 200 patients, but they are transparent about which studies were excluded for this reason (these are listed in Appendix 6, MS p.230) and the ERG agrees that it is reasonable to exclude these studies. Issues of bias and study quality are not considered at the searching, screening or selection stages of the review, but the manufacturer provides a critical appraisal of the included studies in Section 6.4 of the MS (p.76) and Appendix 3 (p.216).

The MS includes a flow diagram that shows the number of publications identified through the database searches and the number of publications included and excluded at each stage of the review (MS p.32). Reasons for excluding studies at the full publication review stage, along with the number excluded for each reason, are detailed in the diagram. Additionally, in the

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

appendices, the manufacturer has provided lists of the RCT and non-RCT studies identified through the searches, and, where relevant, has recorded reasons for excluding publications in these lists (MS p.215 and p.232).

3.1.3 Identified studies

The MS identified two phase III RCTs,[1;2] shown in Table 1, published in eight publications (one journal article and three conference abstracts for each RCT).[1-8] The MS did not identify any relevant non-RCT studies. Both the identified RCTs were sponsored by the manufacturer in collaboration with other organisations. The GOG-0218 trial was sponsored by Genentech (part of the Roche Group) and the National Cancer Institute, and the ICON7 trial was sponsored by Roche, the National Institute for Health Research (through the National Cancer Research Network) and the Medical Research Council. The MS appears to have included all relevant RCTs. The ERG searches did not identify any other relevant studies.

The manufacturer did not supply copies of any of the eight publications identified to the ERG, but the ERG was able to access the original journal articles for each RCT. The manufacturer provided electronic copies of other references cited in the submission.

Table 1: List of included studies

Trial	Patient population	Intervention 1	Intervention 2	Comparator
GOG-0218[1;2]	Epithelial ovarian cancer, primary peritoneal cancer and fallopian tube cancer with stage III optimal (macroscopic), stage III sub-optimal or stage IV disease	CPB15	CPB15+	CPP
		Carboplatin (AUC6) and paclitaxel (175mg/m^2) (q3w 6 cycles) with concurrent IV bevacizumab (15 mg/kg) (q3w for 5 cycles), followed by placebo (q3w for 16 cycles)	Carboplatin (AUC6) and paclitaxel (175mg/m^2) (q3w 6 cycles) with concurrent IV bevacizumab (15 mg/kg) (q3w for 5 cycles), followed by extended bevacizumab (15mg/kg q3w for 16 cycles)	Carboplatin (AUC6) and paclitaxel (175mg/m^2) (q3w 6 cycles), and placebo (q3w for 5 cycles), followed by placebo (q3w for 16 cycles)
ICON7[1;2]	Epithelial ovarian carcinoma,	CPB7.5+	-	CP

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

	primary peritoneal carcinoma or fallopian tube carcinoma with high risk early stage (FIGO stage I/IIA clear cell or grade 3 carcinoma) or advanced stage (FIGO stage IIB-IV, all grades and all histological subtypes) disease	Carboplatin (AUC5 or 6) and paclitaxel (175mg/m^2) with concurrent bevacizumab (7.5mg/kg) (q3w for 6 cycles), followed by extended bevacizumab (7.5mg/kg q3w for 12 cycles)	-	Carboplatin (AUC5 or 6) and paclitaxel (175mg/m^2) (q3w for 6 cycles)

The GOG-0218 trial meets the inclusion criteria of the systematic review; however, the ICON7 trial does not as it included patients with high risk early stage ovarian cancer in addition to patients with advanced ovarian cancer. The ICON7 trial only matches the patient population specified in the scope through a subgroup analysis of patients with advanced disease. It also used a dose of 7.5 mg/kg which is unlicensed. As this study does not match the scope or the licensed indication, we have restricted our review of the evidence submitted by the manufacturer to the GOG-0218 trial in the main part of this report; however, we have added commentary on the ICON7 trial to Appendix 1 for information and as it was used as supporting evidence for the marketing authorisation.

The MS provides an overview of the GOG-0218 trial intervention, comparators, population and number of patients included in the intention to treat (ITT) analyses (MS Table 3, p.33). Further information about the trial design, intervention, comparator, and number of patients randomised is also provided (MS Table 4, p.37 to 42), along with details about the location of the study, the primary and secondary outcomes, and the duration of follow-up. The trial design, intervention and comparator are also shown graphically (MS Figure 2, p.43). Further information about the outcomes is provided in the text on MS p.59-61. Patient inclusion and exclusion criteria are shown (MS Table 5, p.46-47). The number of patients randomised and allocated to each trial arm is shown in a flow diagram (MS Figures 4, p.75). The number of patients screened for eligibility is not reported, but information in the original paper[1;2] shows that all patients enrolled were randomised. The flow diagram also details the number of patients who dropped out and reasons for this. Patients in the placebo arm were allowed to cross-over to receive bevacizumab, but the MS does not state the number of patients who crossed-over in the flow

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

chart. The manufacturer states later on in the MS that "the most recent data cut suggests that 40% of the CPP patients in this study have now received bevacizumab in their subsequent therapy" (MS p.60). Statistical analysis information is provided in the text on MS p.62 to p.69, including power/sample size calculations, hypotheses, statistical test methods, definitions of the analysis populations (including the ITT population), and subgroup analyses. However, this information was incomplete and the manufacturer was contacted for clarification.

The MS states that baseline patient characteristics were similar across treatment arms within both trials and the ERG agrees with this conclusion.

In Section 1.6 of the MS (p.12), it is stated that there are no ongoing trials that are likely to provide additional evidence in the next 12 months. The ERG agrees with this and did not identify any other ongoing relevant trials.

3.1.4 Description and critique of the approach to validity assessment

The MS provides a quality assessment for both RCTs (Appendix 3, MS p.217) and a summary assessment for each RCT is tabulated in Section 6.4 of the MS (p.78). The quality assessment in the MS follows the NICE criteria and is appropriate. Table 2 shows the ERG independent assessment of study quality for the GOG-0218 trial and the MS assessment. As this table shows, the ERG assessment partly agrees with that of the manufacturer.

Table 2: Manufacturer and ERG assessment of trial quality for the GOG-0218 trial

NICE QA Criteria for RCT	MS response	ERG response
1. Was the method used to generate random allocations adequate?	Yes	Yes
2. Was the allocation adequately concealed?	Yes	Yes
3. Were the groups similar at the outset of the study in terms of prognostic factors, e.g. severity of disease?	Yes	Yes
4. 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)?	Yes	No. A change in protocol part way through the trial meant that patients and investigators could be unblinded to treatment allocations once patients had experienced disease progression.[2]
5. Were there any unexpected imbalances in drop-outs between groups? If so, were	No	Yes. Proportionally fewer patients withdrew from

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

they explained or adjusted for?		treatment in the CPB15+ arm than in the CPB15 and CPP arms. This was because fewer patients in the CPB15+ arm withdrew due to disease progression than in the other arms. This was adjusted for by use of ITT analysis for the PFS and OS outcomes.
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
7. Did the analysis include an intention to treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes. PFS (the primary outcome) and OS analyses were conducted using ITT analysis.

3.1.5 Description and critique of manufacturer's outcome selection

All the outcomes reported in the MS by the manufacturer for the GOG-0218 study are appropriate and match the scope. The MS reports all relevant outcomes that are in the GOG-0218 trial publication.[2] The primary outcome was progression-free survival (PFS). This was defined as the period from randomisation to disease progression or death, and was measured by investigator assessment using any of the following measures: global clinical deterioration, CA-125 progression and RECIST criteria. The MS presents three separate analyses of PFS:

- Investigator assessment of PFS censored for CA-125 progression
- Investigator assessment of PFS uncensored for CA-125 progression (a sensitivity analysis; protocol-specified)
- Independent Review Committee (IRC) assessment of PFS censored for CA-125 progression (a sensitivity analysis)

For the primary Regulatory analysis, which is presented in the MS, PFS was censored for progression based on the CA-125 criteria alone. The MS suggests that it is reasonable to censor for CA-125 as it is not a reliable measure of disease progression. The ERG notes, however, that CA-125 measurement is commonly used in the UK for disease progression. Therefore, results not censoring for increased CA-125 are of more relevance to the UK. The ERG also suggests that the uncensored analysis of PFS is more robust and less likely to provide a biased estimate than the censored analyses. The scope specifies that if evidence allows, subgroup analyses of patients with optimally and sub-optimally debulked disease should be presented. The manufacturer has provided exploratory subgroup analyses of PFS by

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

disease stage and debulking status in Table 14 (MS p.84) and by various baseline risk factors in Table 13 (MS p.83).

Secondary outcomes were:

- Overall survival (OS)
- Objective response rate (ORR)
- Adverse events (AEs)
- Health-related quality of life (HRQoL)

The ERG notes that the original paper[2] states that OS was the primary outcome specified in the protocol, but this was changed part the way through the trial to PFS. This was changed as the maintenance of blinding of treatment assignments following disease progression required "to protect the integrity of the data on overall survival" (Burger and colleagues[2] p. 2482) was contested by patients and investigators.

AEs are reported in the main AE section of the MS (p.101 to p.114).

HRQoL was assessed by the Trial Outcome Index of the Functional Assessment of Cancer Therapy-Ovary Trial Outcome Index (FACT-O-TOI) survey, which is a patient self-report measure. The manufacturer presents the HRQoL findings in the MS for the following sub-scales derived from this measure:

- TOI (the primary outcome measure)
- Ovarian Cancer Subscale (OCS) (which also forms part of the TOI score)
- Abdominal discomfort score (ADS)

The MS states that this measure is appropriate for use in oncology clinical trials and in clinical practice, but the MS does not provide a reference to the original source of this scale or information about its reliability and validity in either this or other research. (The ERG requested clarification from the manufacturer about this and the manufacturer provided references that confirm it is a reliable and valid measure.[9-11])

3.1.6 Description and critique of the manufacturer's approach to trial statistics

The MS reports the trial results for all relevant outcome measures. PFS in each bevacizumab arm (CPB15 and CPB15+) was compared to PFS in the CPP arm (referred to as the 'initial

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

primary efficacy analysis' in the MS). If PFS was found to be better in both the bevacizumab-containing arms, they were then compared against each other (referred to as the 'late primary analysis' in the MS). The manufacturer used one-tailed tests in the initial and late primary analyses and took a conservative approach to making statistical comparisons by applying a Dunnett procedure to adjust for multiple comparisons in the initial primary analysis. PFS was analysed using a stratified log-rank test, with initial GOG performance status and disease stage as stratification factors. The MS also reports results from an unstratified log rank test in which median PFS for each treatment arm was estimated using Kaplan-Meier methodology. Along with the median PFS for each group, the MS presents the number of events (but only for some of the analyses), stratified hazard ratios and associated 95% CIs, and p-values. The same approach to data analysis was used in the OS analyses as for the PFS analyses, except that the two bevacizumab arms were not compared. ORR in each bevacizumab arm was compared to the CPP arm using the Cochran-Mantel-Haenszel test, stratified for the same factors as used in the PFS analysis. The number and percentage of patients with an objective response is presented, along with the associated p-values, for the investigator and IRC assessments. The MS also presents the number and percentage of patients who had a complete response, partial response, stable disease or progressive disease, as well as the number and percentage of patients in whom evaluation was not possible. The number of patients included in each analysis is provided for all analyses.

For the subgroup analyses of PFS, the results for the CPB15 and CPB15+ arms were pooled for all analyses except the analysis of PFS by patients' disease stage and debulking status. The MS states that results were pooled as these two arms received identical treatment prior to Cycle 7. The ERG notes that these two arms were not pooled for the subgroup analyses reported in the original trial paper.[2] It is not clear if the decision to pool arms was planned or made post-hoc, whether interim or final data are presented (the data cut-off date for this analysis is not provided) and up to which treatment or follow-up timepoint in the study PFS is reported for the subgroup analyses. It is also not clear in the MS whether the subgroup analyses were planned or post-hoc analyses. (Clarification was sought from the manufacturer which confirmed that the majority of the subgroup analyses were planned, except for the analyses by race, baseline sum of longest diameter [SLD] and baseline CA-125.) The ERG notes, though, that the analysis of PFS by patient age uses slightly different age categories to those proposed in the protocol. It remains unclear whether or not the decision to pool the CPB15 and CPB15+ arms was planned. Timepoints for analyses are shown in Table 3 below. Power calculations were not performed for

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

each of the subgroups, but as the number in each subgroup looks reasonable and the associated confidence intervals are narrow, this may not be an issue of particular concern.

The OCS and ADS HRQoL data were presented in the MS using a graphical display of mean change in scores over time from baseline for each treatment arm. In terms of the TOI HRQoL, change from baseline within groups was assessed in terms of whether or not the difference in the mean score was clinically meaningful (defined as a mean improvement of at least 5 points, based on guidance for FACT-O-TOI). In addition to changes in HRQoL within groups, the MS states that three hypotheses that changes in HRQoL were independent of the treatment received were tested, using a mixed effect model. To do this, the following between group comparisons of TOI scores were tested:

- CPP versus the pooled bevacizumab arms prior to Cycles 4 and 7
- CPP versus CPB15+ arm prior to Cycles 4 to 21
- CPB15 versus CPB15+ arms prior to Cycles 13 and 21

The ERG has noted the following issues with the HRQoL analyses:

- The MS states that the above three hypotheses were based on the "protocol specifications with modifications" (MS p.68). It is not clear how these analyses deviate from those specified in the protocol. (The ERG requested clarification from the manufacturer, and the manufacturer was unable to confirm what modifications had been made).
- The MS states that each hypothesis was tested using a two-sided alpha level of 0.05, but the original paper[2] states that hypotheses were tested using an alpha level of 0.0167 (as a result of adjusting for multiple comparisons). It is not clear why a less conservative alpha level was used in the MS or what approach was specified in the protocol.
- The number of patients in each analysis and p-values for differences between groups are not reported in the results section; findings are only reported in terms of mean score changes with some commentary on statistically significant changes. The ERG requested p-values for these analyses from the manufacturer, but the manufacturer was unable to supply these.
- The MS notes that one of the ovarian cancer subscale items of the FACT-O-TOI was omitted from the overall total score for the analyses reported in the MS, but this is not stated in the trial paper.[2] It is unclear if omission of this item was planned or made post-hoc.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

- The text summary of change in OCS scores over time (MS p.87) makes reference to whether or not changes were clinically meaningful (defined as a difference of ≥ 3 points, p 86), but the manufacturer has not provided a justification for this.

The PFS and OS analyses were conducted using ITT analyses. ITT analyses were not used to assess ORR or AEs. Analysis of ORR was only carried out with patients who had measurable disease at baseline. For AEs, the manufacturer presents safety data for patients who had received study treatment at least at Cycle 2 or beyond. As the difference between the number of patients randomised and the number included in the AE analyses is small (n ≤ 24 in each arm) this is unlikely to have affected the results. It is unclear if ITT analyses were used to analyse the HRQoL data.

There is inconsistency in the dates of the analyses reported in the MS, so it is not clear if interim or final data have been reported for some analyses. (Clarification sought from the manufacturer reports that the exploratory subgroup analysis by disease stage and debulking status was not subsequently updated using data as of September 2010; the manufacturer appears to have provided the most up-to-date data available for most outcomes, see Table 3.)

Table 3: Dates for analyses presented in the MS for PFS and OS

Outcome measure	**Primary analysis For FDA February 2010**	**Final Analysis Cut-off date September 2010**	**Updated analysis Cut-off date August 2011**
PFS censored for CA-125		Investigator Assessed (MS, Table 10) IRC (clarifications, A3)	n/a
PFS not censored for CA-125 ('per protocol')	Investigator Assessed (MS, Table 12)		Investigator Assessed (MS, p.82 text)
PFS Exploratory Subgroup	Investigator Assessed - disease stage and debulking status (MS, Table 14 and clarifications, A3)	Investigator Assessed - baseline risk factors (MS, Table 13 and clarifications, A3)	n/a
OS			Final (MS, Table 15))

n/a not available

Overall, the manufacturer's approach to data analysis is appropriate, but the subgroup analyses using the pooled CPB15 and CPB15+ arms and the HRQoL analyses should be interpreted with caution.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

3.1.7 Description and critique of the manufacturer's approach to the evidence synthesis

The MS provides a narrative synthesis of the findings of the GOG-0218 trial. Some of the tabulated and narrative data in the MS differ to that reported in the trial paper for the OS, HRQoL analyses and subgroup analyses of PFS.[2] In terms of differences in the OS and HRQoL data reported in the MS and the trial paper,[2] these do not change the interpretation of the data. The use of different alpha levels in the MS and trial paper for detecting statistically significant differences between the arms in the HRQoL analysis also do not affect the conclusions made about group differences for the HRQoL outcome.

In terms of the subgroup analyses of PFS, the analyses comparing the pooled bevacizumab-containing arms with the CPP arm are not comparable to the subgroup analyses reported in the original trial paper[2] as in the paper the bevacizumab-containing arms were not pooled. The ERG notes that when the bevacizumab-containing arms were pooled, they generally showed favourable effects over the CPP arm across subgroups; however, in the original paper,[2] the subgroup analyses showed that while the CPB15+ arm was generally superior in effectiveness across subgroups to the CPP arm, the CPB15 arm was not. The data for the subgroup analysis of PFS by disease stage and debulking status reported in the MS for each of the three arms differ to that reported in the trial paper,[2] but the differences are minor and do not affect the interpretation of the data.

A meta-analysis of the GOG-0218 and ICON7 trials is not provided. The MS states that this is because the trials are not comparable in terms of treatment dose and duration and patient population. The ERG agrees with this decision.

3.2 Summary statement of manufacturer's approach

The quality of the MS based on CRD criteria[12] for a systematic review as assessed by the ERG is shown in Table 4.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Table 4: Quality assessment (CRD criteria) of MS review

CRD Quality Item: score Yes/ No/ Uncertain with comments	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes. Inclusion and exclusion criteria are reported (details given in MS Appendix 2). However, the inclusion criteria are wider than the scope/decision problem in that patients with any stage of ovarian cancer are included, rather than those with Stage III or IV only.
2. Is there evidence of a substantial effort to search for all relevant research? Ie all studies identified	Yes. Extensive searches were conducted for clinical and cost effectiveness.
3. Is the validity of included studies adequately assessed?	Yes. The validity of the included RCTs was undertaken using standard CRD criteria for assessing the quality of RCTs and is presented in a summary table and appendix only (MS Table 9, Appendix 3). No narrative discussion is presented.
4. Is sufficient detail of the individual studies presented?	Yes. Study characteristics are described for the included RCTs and presented in several tables.
5. Are the primary studies summarised appropriately?	Yes. The primary studies are appropriately summarised. The key RCT is summarised through narrative means and tabulation of results for investigator assessed and IRC analyses for PFS for the ITT population. Other outcomes are appropriately summarised. Summary similarities and differences between the primary trial and the additional supporting evidence are mentioned.

The systematic review is of reasonable quality according to CRD criteria and the submitted evidence reflects the decision problem defined in the MS although no details are given for any of the processes used in the systematic review (ie whether assessment was by a single reviewer or independently by two reviewers).

Overall the risk of systematic error in the systematic review appears to be low. However, as mentioned above, the systematic review was wider than the decision problem/scope and included an additional study which does not meet the population defined in the scope.

3.3 Summary of submitted evidence

In this section of the report the ERG provides a summary of the evidence presented in the MS from the key RCT (GOG-0218).[2] Results from the additional trial (ICON7[1]) which did not match the scope or the MS review inclusion criteria, are available in Appendix 1 of this report. Data have been checked by the ERG and summarised for the primary outcome and key secondary

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

outcomes. Some points of clarification were requested from the manufacturer and these are noted below.

Summary of results for progression free survival (GOG-0218)

The primary outcome was PFS (defined as the period from randomisation to disease progression or death) based on investigator assessment censored for CA-125 and non-protocol specified cancer therapy (NPT) prior to disease progression for the ITT population. There was a statistically significant improvement in the median PFS of 6 months in the CPB15+ arm compared to the CPP arm (CPP 12 months, CPB15+ 18 months; HR 0.645, 95%CI 0.551, 0.756, p<0.001). No difference in PFS was found between the CPP arm and CPB15 arm. (See MS Table 10, p.80).

Results derived from the IRC sensitivity analysis are also presented and also show median PFS statistically significantly improved in the CPB15+ arm compared with the CPP arm (CPP 13.1 months, CPB15+ 19.1 months; HR 0.62, 95%CI 0.50, 0.77, p<0.0001). (See MS Table 11, p.81).

Sensitivity analysis using investigator assessed PFS without censoring for CA-125 progression or NPT prior to disease progression, shows similar results although the median PFS for the CPP and CP15+ arms (10.3 and 14.1 months respectively) are shorter than those reported for the censored data. (See MS Table 12, p.82). PFS results presented in the MS are summarised in Table 5. (Clarification requested from the manufacturer on missing data for the updated analyses stated that these are not available).

Table 5: Median PFS for the different analyses from Study GOG-0218 (ITT population)

		CPP (n=625)	CPB15 (n=625)	CPB15+ (n=623)
Investigator assessed[1] (censored)	Median PFS mths	12.0	12.7	18
	Stratified hazard ratio (95% CI)		0.84 (0.71, 0.99)	0.645 (0.551, 0.756)
	P value		0.204	<0.001
Independent review Committee[1] (censored)	Median PFS mths	13.1	13.2	19.1
	Stratified hazard ratio (95% CI)		0.93 (0.76, 1.13)	0.62 (0.50, 0.77)
	P value		0.222	<0.0001
Investigator assessed[2]	Median PFS mths	10.3	11.2	14.1
	Stratified hazard		0.908	0.717

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

(not censored)	ratio (95% CI)		(0.795, 1.040)	(0.625, 0.824)
	P value		0.16	<0.0001
Investigator assessed[3] (not censored) Updated	Median PFS mths	n/a	n/a	n/a
	Stratified hazard ratio (95% CI)		n/a	0.77 (0.681, 0.870)
	P value			n/a

[1]Final analysis, September 2010; [2] Primary analysis, February 2010; [3] Updated analysis, August 2011; n/a not available

The ERG is advised by a clinical expert that the updated investigator assessed (not censored) HR (0.77) is the most appropriate for the UK.

Summary of results for overall survival (GOG-0218)

The final OS analysis of the GOG-0218 study is reported (MS Table 15, p.85) and shows a non-significant increase in median OS of 3.2 months in the CPB15+ arm compared with the CPP arm (CPP 40.6 months, CPB15+ 43.8 months; HR 0.88, 95%CI 0.75, 10.4, p=0.0641). OS results are shown here in Table 6.

Table 6: Overall survival for Study GOG-0218 (ITT population)

		CPP (n=625)	**CPB15 (n=625)**	**CPB15+ (n=623)**
GOG-0218	Median mths	40.6	38.8	43.8
	Hazard ratio (95% CI)		1.07 (0.91, 1.25)	0.88 (0.75, 1.04)
	P value		0.2197	0.0641

Summary of results for objective response rate (GOG-0218)

The ORR according to investigator assessment in Study GOG-0218 shows a non-significant increase of 2.6% in the CPB15+ arm (66%) compared with the CPP arm (63.4%); however, the IRC analysis reports a significant increase of 8.6% in the CPB15+ arm compared with the CPP arm (77.4% vs 68.8%, p<0.0012). The IRC analysis also reports a statistically significant difference between the CPB15 and CPP arms (75.4% vs 68.8%, p<0.0106) (See MS Table 16, p.85).

Summary of Health related quality of life (GOG-0218)

HRQoL was also reported. Although bevacizumab-containing therapy produced some QoL disruptions during chemotherapy, differences between treatment arms were small and not clinically meaningful.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Sub-group analyses results (GOG-0218)

Exploratory PFS analyses for various subgroups using baseline risk factors are also reported (MS Table 13, p.83). Results for these subgroups are consistent with the primary analysis with hazard ratios less than one, favouring the combined CPB15+ and CPB15 arms compared with the CPP arm; the difference in the median PFS is close to the 6 month benefit reported in the primary analysis. Subgroup analysis by disease stage and debulking status are presented (MS Table 14, p.84) which show a statistically significant increase in PFS for all stages of disease with CPB15+ compared with CPP alone. No significant increase was shown for the CPB15 arm. PFS results for these subgroup analyses are shown here in Table 7.

(Clarification from the manufacturer was requested on missing p values for subgroup analyses as p values were given for other analyses; the response stated that p values were not available for these evaluations).

Table 7: Median PFS by disease stage and debulking status from Study GOG-0218 (uncensored data)

		CPP (n=219)	CPB15 (n=204)	CPB15+ (n=216)
Stage III optimally debulked disease	Median PFS mths	12.4	14.3	17.5
	Stratified hazard ratio (95% CI)		0.81 (0.62, 1.05)	0.66 (0.50, 0.86)
	P value		n/a	n/a
Stage III Sub-optimally debulked disease	Median PFS mths	10.1	10.9	13.9
	Stratified hazard ratio (95% CI)		0.93 (0.77, 1.44)	0.78 (0.63, 0.96)
	P value		n/a	n/a
Stage IV	Median PFS mths	9.5	10.4	12.8
	Stratified hazard ratio (95% CI)		0.90 (0.70, 1.16)	0.64 (0.49, 0.82)
	P value		n/a	n/a

n/a not available

Summary of adverse events (GOG-0218)

The MS reports on the safety of bevacizumab combined with carboplatin and paclitaxel in comparison to chemotherapy using carboplatin with paclitaxel alone. It states that as AEs were secondary outcomes in the GOG-0218 trial (and ICON7), no additional searches were carried out to identify studies that included AEs.

The MS provides a summary overview of AEs, including any AE experienced and death (MS, Table 24, p.103), and reports AEs of all grades that occurred with a ≥ 5% difference between

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

groups (MS, Table 25, p.104). It is not clear why only AEs with a ≥ 5% difference between groups are reported (clarification was sought from the manufacturer, which stated that this difference level was chosen "to demonstrate where the major differences lie" in safety between arms). However, the MS provides information on AEs of all grades of special interest to bevacizumab treatment (MS, Table 26, p.105). The MS reports only AEs that occurred between patients starting Cycle 2 of treatment to 30 days after patients received their last study treatment, as the study drug was not started until Cycle 2. Incidence of AEs is reported for each arm as the number and percentage of patients experiencing each event. The MS does not provide p-values nor relative risk, risk difference or associated 95% CIs statistics for the analyses (the ERG requested these from the manufacturer, which stated that it was not possible to supply these as the data were not analysed in this way), so it is not possible to tell whether the differences reported are statistically significant. Instead, the manufacturer reports that the incidence of AEs was higher in the bevacizumab-containing groups when the difference in incidence was ≥ 10% higher for the AEs with a ≥ 5% difference between groups. For AEs of special interest to bevacizumab, the manufacturer states that AEs differed between treatment arms when the incidence of AEs was >1% different between groups. (The ERG requested clarification from the manufacturer about whether these were the only statistically significant differences between groups. The manufacturer stated that statistical significance tests were not conducted for these analyses and stated that the criteria they used were chosen as an "arbitrary threshold" to highlight the main differences between the arms to the reader. Based on this, the ERG suggests that the manufacturer's textual summary of differences in AEs between arms should be interpreted with caution.)

Table 8 provides a summary of the AEs that the MS noted as differing between treatment arms, based on their criteria above. The MS states that proportionally more patients in the bevacizumab-containing arms experienced a grade 3-5 AE than patients in the CPP arm. The death rate did not appear to differ between treatment arms (24.1%, 24.4% and 21.5% of patients died in the CPP, CPB15 and CPB15+ arms, respectively). However, the MS notes on p.112 that "more deaths from AEs were observed in the two bevacizumab-containing arms (9 patients [1.5%] and 14 patients [2.3%] in the CPB15 and CPB15+ arms, respectively) compared with the control arm (4 patients [0.7%] in the CPP arm)". (The ERG requested clarification from the manufacturer about whether or not this was a statistically significant difference. The manufacturer stated that statistical significance tests were not conducted for these analyses.) The ERG notes that discontinuation of study treatment due to AEs appeared to be higher in the

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

bevacizumab-containing arms than the CPP arm (16.4%, 13.7% and 9.7% of patients in the CPB15+, CPB15 and CPP arms, respectively). The MS states that more patients in the bevacizumab-containing arms than the CPP arms experienced the following AEs: stomatitis (all grades), dysarthia (all grades), headache (all grades), epistaxis (all grades), gastrointestinal perforation (grade 3-5), non-CNS bleeding (grade 3-5) and hypertension (all grades and grade 3-5). Using the criteria presented in the MS to determine group differences, the ERG notes that the following AEs of special interest to bevacizumab also differed between the CPP and bevacizumab-containing treatment arms, but this is not commented on in the MS: non-CNS bleeding (all grades), proteinuria (all grades), neutropenia (all grades), and gastrointestinal perforation (all grades).

Table 8: AEs highlighted in the MS as differing between treatment arms (GOG-0218)

Adverse event	**CPP (n = 601)**	**CPB15 (n = 607)**	**CPB15+ (n = 608)**
Grade 3-5 adverse events, excluding laboratory data	274 (45.6%)	307 (50.6%)	337 (55.4%)
Stomatitis (all grades)	80 (13.3%)	117 (19.3%)	147 (24.2%)
Dysarthria (all grades)	9 (1.5%)	58 (9.6%)	72 (11.8%)
Headache (all grades)	126 (21.0%)	156 (25.7%)	202 (33.2%)
Epistaxis (all grades)	55 (9.2%)	182 (30.0%)	184 (30.3%)
Hypertension (all grades)	81 (13.5%)	143 (23.6%)	196 (32.2%)
Hypertension (grade 3-5)	12 (2.0%)	34 (5.6%)	60 (9.9%)
Gastrointestinal perforation (grade 3-5)	2 (0.3%)	10 (1.6%)	10 (1.6%)
Non-CNS bleeding (grade 3-5)	5 (0.8%)	8 (1.3%)	13 (2.1%)

Note. The MS does not provide information about whether or not these were statistically significant differences or whether any other AE showed statistically significant differences between groups.

The MS also presents a secondary analysis of AEs, that compares the number of AEs reported during the chemotherapy (Cycles 2-6) and maintenance phases of treatment (Cycles 7-22) in each trial arm. AEs that occurred with an incidence rate difference of ≥ 5% between groups are presented in Tables 28 and 29 in the MS (p.108). For the incidence of AEs during chemotherapy, the MS presents a pooled incidence rate for the CPB15 and CPB15+ arms and compares this with the incidence rate for the CPP arm. These tables show that the following AEs were more common in the CPB15/CPB15+ arms than the CPP arm during chemotherapy: stomatitis, dysarthria, headache, dyspnea, epistaxis and hypertension. The data for the maintenance phase show that the incidence of the following AEs in the CPB15+ arm was higher than in the CPP arm: diarrhoea, arthralgia, myalgia, pain in extremity, dysarthria, headache, epistaxis and hypertension.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

A clinical expert consulted by the ERG indicated that the AEs associated with bevacizumab, including hypertension, are manageable in clinical practice. Bowel perforation is a more serious risk, but this most commonly occurs in the later stages of ovarian cancer in conjunction with other bowel problems.

3.4 Summary

Results of the key Phase III RCT show that patients with advanced ovarian cancer who received first-line therapy with bevacizumab in combination with chemotherapy (carboplatin and paclitaxel) followed by bevacizumab maintenance alone had statistically significant and clinically meaningful improvement in PFS compared with patients who received chemotherapy plus placebo. Patients who received bevacizumab in combination with chemotherapy without bevacizumab maintenance therapy did not have a statistically significantly improved PFS compared with patients who received chemotherapy alone.

Sensitivity analyses (using IRC review and PFS not censored for CA-125) suggest robustness of the primary results. Also the observed PFS benefit in the CPB15+ group was shown for exploratory subgroup analysis by disease stage and debulking status.

No difference in OS between the CPB15+ arm and the CPP arm was found although results may have been confounded by post-progression cross-over of placebo patients to receive bevacizumab.

AEs for which the incidence was ≥ 10% higher in the bevacizumab-containing arms than the chemotherapy-alone arm were stomatitis, dysarthia, headache, epistaxis, and hypertension.

On the whole it appears that the MS contains an unbiased estimate of treatment effect within the stated scope of the decision problem. In general the manufacturer's interpretation of the evidence is appropriate and justified. It discusses the relevance of the evidence base to UK practice and its limitations. However, some concerns/uncertainties include:

- The different assessments (investigator assessed, IRC, censored, not censored) are not consistently reported for all time points. This suggests that there may have been selective reporting of data and it is not clear what impact this may have on conclusions. (Clarification requested from the manufacturer states that updated PFS data censored

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

for CA-125 are not available; also exploratory analyses were not updated as they were intended only to confirm the validity of investigator assessed PFS). Although the direction of evidence is consistent, the size of effect varies with the different analyses and over time. The hazard ratio presented for PFS ranges from 0.62 (IRC assessed, data censored) to 0.77 (updated investigator assessed, without data censoring, most relevant to UK); this corresponds to a reduction in the risk of progression or death ranging from 23% to 38%. There was between 4 and 6 months gain in median PFS depending on the analysis (18 months in the CPB15+ arm compared with 12 months in the CPP arm for censored data, and 14 months compared with 10 months respectively for non-censored data).

- It is not clear what impact censoring data for increased CA-125 has on the results, although it appears that not censoring for CA-125 gives a more conservative estimate of effectiveness than censoring for CA-125. It is reported in the MS that the majority of UK physicians will use the RECIST criteria guidelines only (MS p.11); however, expert advice to the ERG suggests otherwise and that CA-125 is used routinely in clinical practice. Therefore, results not censoring for increased CA-125 are of more relevance to the UK.
- The licensed dose of 15mg/kg bevacizumab for 15 months was used in the GOG-0218 trial but the dose most likely used in current clinical practice is 7.5 mg/kg;
- Clinical advice to the ERG is that bevacizumab is given to patients with Stage III residual disease, the aim being to maintain PFS for as long as possible so that therapy may be repeated at relapse. However, the GOG-0218 trial included patients with Stage III optimally debulked disease so it may not be entirely applicable to UK practice.
- Although the ICON7 study has been presented as supporting evidence, and more closely reflects current clinical practice in terms of dosage of bevacizumab, the patient groups for the trial and also the subgroup analyses do not completely match the decision problem.
- It should be noted that the MS does not report p-values in some analyses, making interpretation of findings difficult as it is not possible to tell the level at which differences between groups are statistically significant. (Clarification requested from the manufacturer stated that outstanding p values were not available.)

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

4 ECONOMIC EVALUATION

4.1 Overview of manufacturer's economic evaluation

The manufacturer's submission to NICE includes:

i) a review of published economic evaluations of bevacizumab in advanced or metastatic ovarian cancer.

ii) a report of an economic evaluation undertaken for the NICE STA process; the cost effectiveness of bevacizumab in combination with carboplatin and paclitaxel compared with carboplatin and paclitaxel for first-line treatment in women with newly diagnosed stage III or IV ovarian cancer.

Manufacturer's review of published economic evaluations

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of bevacizumab in advanced or metastatic ovarian cancer from a UK perspective using several health economic databases and medical databases (MS p.124). The inclusion and exclusion criteria for the systematic review are listed (in section 7.1 of the MS, p.124). The inclusion criteria state that cost effectiveness studies of bevacizumab in advanced ovarian cancer would be included.

Nine studies were identified and of these six studies were excluded, mainly as they were not cost effectiveness studies. Three studies were included for full review.[13-15] The studies were quality assessed using the Drummond and Jefferson checklist[16] (suggested by NICE) (MS Appendix 11). No interpretation or conclusions of this quality assessment were provided in the MS. Results were presented from the three studies but no discussion or conclusions were given on these results by the manufacturer. The ERG suggests that the review of the published economic evaluations could have been more informative by comparing and discussing the alternative model structures and the corresponding differences in model results from that developed by the manufacturer.

CEA Methods

The cost effectiveness analysis (CEA) uses a 3 state semi-Markov model to estimate the cost-effectiveness of bevacizumab in combination with carboplatin and paclitaxel (CPB) compared with carboplatin and paclitaxel (CP) for first-line treatment in women with newly diagnosed stage

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

III or IV ovarian cancer. The model has health states for PFS, progression and death. The model adopted a time horizon of ten years with a cycle length of 1 week. The model costs and outcomes were discounted at 3.5% per annum. The perspective of the model is the UK NHS Personal Social Services (PSS) and results are presented as incremental cost per QALY gained.

The transition of patients from PFS to disease progression and death were derived from the GOG-0218 (and ICON7) RCTs. HRQoL is used in the model for the health states for PFS and progression, based upon EQ-5D surveys of patients in the ICON7 study. Resource use in each health state was based on a previous NICE appraisal in ovarian cancer[17] and costs were taken from the BNF63,[18] PSSRU[19] and NHS reference costs 2010/11.[20]

Sensitivity analyses were presented for a limited number of parameters (MS section 7.7.7, p186), for PFS extrapolation, post-progression survival, utility values, administration, AEs and chemotherapy costs, time horizon and discounting rates. Probabilistic sensitivity analyses were also presented (MS section 7.7.8, p.188).

The model was validated by comparing OS predicted by the model with an external publication using ovarian cancer patients with similar disease severity and surgical outcomes.

CEA Results

The results from the economic evaluation are presented, based on the GOG-0218 (and ICON7) RCTs. Results are presented as the incremental cost per QALY gained for CPB vs. CP. Base case results are presented (MS Section 7.7, Table B13-B14, p.185) for GOG-0218 (and ICON7).

For the base case an incremental cost per QALY gained of £144,066 is reported for GOG-0218 (and £31,592 for ICON7) as shown in Table 9. The key drivers of the cost effectiveness results are the dose and cost of bevacizumab and the duration of the treatment (MS p.192).

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Table 9: Base case cost effectiveness results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus incremental (LYG)	ICER (£) incremental (QALYs)
GOG-0218								
CP	£17,166	3.985	2.973					
CPB	£44,254	4.212	3.161	£27,089	0.228	0.188	£118,876	£144,066
ICON7								
CP	£16,111	3.066	2.278					
CPB	£33,841	3.809	2.839	£17,729	0.743	0.561	£23,846	£31,592

The MS summarises the results of the PSA stating that there is a 0% probability of bevacizumab being cost-effective, relative to treatment with carboplatin and paclitaxel only, at a threshold willingness to pay of £30,000 per QALY gained, based upon the GOG-0218 RCT (and 42% based upon the ICON7 RCT).

4.2 Critical appraisal of the manufacturer's submitted economic evaluation

Critical appraisal of manufacturer's submitted economic evaluation

The ERG has considered the methods applied in the economic evaluation in the context of the critical appraisal questions listed in Table 10 below, drawn from common checklists for economic evaluation methods (e.g. Drummond and colleagues[21]). The critical appraisal checklist indicates that overall the manufacturer follows recommended methodological guidelines, with the exception of the time horizon, where a 10 year time horizon has been used, rather than a lifetime horizon.

Table 10: Critical appraisal checklist of economic evaluation

Item	Critical Appraisal	Reviewer Comment
Is there a well defined question?	Yes	On page 8 the MS states, 'A cost utility analysis was conducted comparing bevacizumab in combination with carboplatin and paclitaxel against carboplatin and paclitaxel chemotherapy alone in ovarian cancer patients with advanced disease'.
Is there a clear description of alternatives?	Yes	Bevacizumab in combination with carboplatin and paclitaxel against carboplatin and paclitaxel chemotherapy
Has the correct patient group / population of interest been clearly stated?	Yes	The submission presents data from GOG-0218 ITT population as the base case. *(Discussed in sections 4.2.2 of this report)*
Is the correct comparator used?	Yes	Carboplatin and paclitaxel *(Discussed in section 4.2.3 of this report)*
Is the study type reasonable?	Yes	Cost utility analysis

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Is the perspective of the analysis clearly stated?	Yes	NHS / PSS
Is the perspective employed appropriate?	Yes	According to the NICE reference case
Is effectiveness of the intervention established?	Yes	Patient level data used from the GOG-0218 trial for PFS and OS. *(Discussed in section 4.2.4 of this report)*
Has a lifetime horizon been used for analysis?	No	A time horizon of 10 years has been used. *(Discussed in section 4.2.1 of this report)*
Are the costs and consequences consistent with the perspective employed?	Yes	*(Discussed in sections 4.2.6 / 4.2.7 for costs and4.2.5 for outcomes of this report)*
Is differential timing considered?	Yes	Costs and health benefits discounted at 3.5% per year.
Is incremental analysis performed?	Yes	Given in MS Section 7.7, Table B13-B14 for the base case results.
Is sensitivity analysis undertaken and presented clearly?	Yes	One way sensitivity analysis is presented in MS Table 63/64.

NICE reference case

The NICE reference case requirements have also been considered for critical appraisal of the submitted economic evaluation in Table 11. The MS analysis follows the NICE reference case.

Table 11: NICE reference case requirements

NICE reference case requirements:	Included in submission	Comment
Decision problem: As per the scope developed by NICE	Yes?	Results also presented for results for an unlicensed dosage of bevacizumab of 7.5 mg/kg of the ICON7 RCT.
Comparator: Alternative therapies routinely used in the UK NHS	Yes	
Perspective on costs: NHS and PSS	Yes	
Perspective on outcomes: All health effects on individuals	Yes	
Type of economic evaluation: Cost effectiveness analysis	Yes	
Synthesis of evidence on outcomes: Based on a systematic review	Yes	
Measure of health benefits: QALYs	Yes	*(Discussed in section 4.2.5 of this report)*
Description of health states for QALY calculations: Use of a standardised and validated generic instrument	Yes	*(Discussed in section 4.2.5 of this report)*
Method of preference elicitation for health state values: Choice based method (e.g. TTO, SG, not rating scale)	Yes	
Source of preference data: Representative sample of	Yes	

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

the public		
Discount rate: 3.5% pa for costs and health effects	Yes	

4.2.1 Modelling approach / Model Structure

The MS cost effectiveness analysis uses a 3-state semi-Markov model to estimate the cost effectiveness of bevacizumab with health states consisting of PFS, Progression and Death. The model was developed in Microsoft Excel. A schematic of the model is shown (MS Figure 15, p.131). The MS states that a Markov model was chosen *'primarily due to the confounding of OS as a consequence of the large proportion of patients randomized to the chemotherapy group who later received bevacizumab following their initial disease progression (approximately 27.7%). As a consequence the analysis was simplified by assuming a similar rate of death post-progression in both arms.'* The ERG notes that an alternative model structure was used for the ICON7 analysis, i.e. a 3-state area-under-the-curve model.

The perspective of the model is the UK NHS PSS and results are presented as the incremental cost per QALY gained. The model adopted a time horizon of ten years with a cycle length of 1 week. The model costs and outcomes were discounted at 3.5% per annum. The MS justified the time horizon by stating that this is the duration of reliable long term survival in the target cohort. The ERG notes that after ten years about 10% of patients are still alive in the model and therefore considers that a longer time horizon should have been adopted.

The MS justifies the model structure and the health states by stating they are typical of modelling in metastatic oncology and similar structures have been utilised in numerous NICE appraisals including those specifically in advanced ovarian cancer.[17] The ERG considers that the structure of the model chosen is suitable for this health condition and the disease states and pathways reflect the underlying biological condition.

In the model, all patients start in the PFS state. Patients move to the progression state according to the PFS trial data from GOG-0218 until 28 months, after which PFS is represented by a log logistic parametric function. After disease progression, patients may progress to death according to a constant probability. The model then allocates health state utility values to patients in the PFS and progression health states.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

The MS describes the following structural assumptions: the base case models assumed that no vial sharing was permitted for patients receiving bevacizumab (although this was tested in the sensitivity analysis) and AEs requiring treatment were assumed to occur in the first week of the model. The ERG notes that bevacizumab is not administered until cycle two, whilst the model includes these AEs during the first cycle. Although this is unlikely to be consistent with clinical practice, the ERG considers that this assumption would have a negligible effect on the model results.

4.2.2 Patient Group

The patient group in the manufacturer's model is not described in detail, although it is implicit that it is similar to the population of the GOG-0218 clinical trial. Model variables are provided (MS Table B2; p.142) only for age, weight, height and body surface area. These are taken from a retrospective study of the characteristics of cancer patients reported by Sacco and colleagues.[22] No explanation is provided for why the patient characteristics used in the model were not taken from the GOG-0218 trial. In the model, patients' age = 56.34 years, weight = 60.49 ± 13.08 kg, height = 161.87 cm and body surface area = 1.71 ± 0.1802 m^2 (no variance estimates are provided for age and height). Although these four population characteristics alone would be consistent with the NICE scope, other key population characteristics for defining the population relevance, such as performance status and disease status according to FIGO staging, are not reported in the MS for the model (although they are reported for the clinical trials).

The MS states that two different analyses were conducted, one based on the GOG-0218 trial which reflects the dose of bevacizumab as specified in the NICE scope (15 mg/kg every 3 weeks), and another analysis based on the ICON7 trial which reflects a lower dose of bevacizumab that is outside the NICE scope (7.5 mg/kg every 3 weeks). It is argued by the manufacturer (and agreed by the clinical expert consulted by the ERG) that the lower dose is representative of current clinical practice.

The MS does not explicitly report population exclusion criteria in relation to the economic model. The inclusion and exclusion criteria for the GOG-0218 and ICON7 trials appear implicitly in the MS to apply to the model analyses. If so, the exclusion criteria listed for the GOG-0218 trial (Table 5 in the MS; p.46-47) may limit the generalisability of the health economic analysis

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

findings to patients who do not have co-morbidities (including other cancers) and who have not had prior surgery.

Numerous subgroup analyses were conducted in GOG-0218 trial (for PFS reported by subgroups see section 3.3) but subgroups were not referred to in the model that was based on this trial.

The ICON7 trial falls outside of the NICE scope since it included patients with less severe initial disease (FIGO classes I and II) than specified in the scope (which is restricted to FIGO classes III and IV). The model based on the ICON7 trial used data only from the expanded high-risk patient subgroup (FIGO assessment criteria III or IV or inoperable patients) (MS section 7.9) as it is stated in the MS that this population would be within the NICE scope. However, the ERG notes that this population may not be fully within the scope because it includes a group of inoperable patients. Furthermore, as noted above, the analysis based on the ICON7 trial remains outside the NICE scope on account of the low bevacizumab dose employed. There do not appear to be key subgroups that are missing from the model analysis. Due to the mode of action of bevacizumab, a subgroup of patients with higher expression of VEGF may be more likely to benefit from bevacizumab treatment but no specific biochemical markers are currently available.

4.2.3 Interventions and comparators

The intervention included in the model is not explicitly described in the MS but can be inferred from Table 1 of the MS (p.13). The intervention was based on that used in the GOG-0218 trial, and comprised six 21-day cycles of dual-agent intravenous chemotherapy (carboplatin and paclitaxel), with intravenous bevacizumab started in the second cycle and then continued for 21 cycles. The intervention thus comprised one cycle of chemotherapy alone, five cycles of chemotherapy + bevacizumab, and 16 cycles of bevacizumab alone. As noted above, the dose of bevacizumab was 15 mg/kg once per 21-day cycle.

The comparator employed in the model is not explicitly described in the MS. It is implied in the model that the comparator reflected best supportive care, which included chemotherapy. The comparator appears to be based on that used in the GOG-0218 trial, in which bevacizumab was replaced by an intravenous placebo. The comparator treatment thus comprised one cycle of chemotherapy alone, five cycles of chemotherapy + placebo, and 16 cycles of placebo alone.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

However, the composition of the placebo is unclear as it not reported by the MS and was not specified in the primary publication (or its supplementary Appendix) for the GOG-0218 trial.[2]

Both the intervention and comparator are consistent with the NICE scope and reflective of current UK practice, with the exception of the dose of bevacizumab employed. As noted above, current clinical practice may use a lower dose of bevacizumab (7.5 mg/kg once per 21-day cycle) than that specified in the NICE scope.

4.2.4 Clinical Effectiveness

The MS reports that the model uses the primary outcome from the GOG-0218 trial, namely PFS. PFS is defined (MS Table 8, p.55) and a rationale for its use is given (MS p.59). The model uses Kaplan-Meier survival curves for PFS until the convergence of survival functions for the intervention and comparator (month 28) (MS section 7.3.1.1, p.136). To extrapolate survival times beyond clinical follow-up, the model utilises a parametric survival model (log-logistic) beyond month 28. Extrapolated data are reported in Figure 16 of the MS (p.136).

The results of the GOG-0218 trial include various Kaplan Meier (KM) curves for PFS and the MS notes that the 'updated PFS analysis' was used in the model (MS section 7.2.2.1, p.132). This analysis included censoring for patients who were presumed to experience progression based on rising CA-125 levels, or who switched to non-protocol therapies. The precise source of these data is not made clear in the MS but they are consistent with a Kaplan Meier survival curve which is given in Figure 3 of a supplementary appendix to the primary publication of the GOG-0218 trial.[2] The MS does not discuss the rationale for choosing these data for PFS and they may not be the most robust data for the appraisal. The censoring for rising CA-125 means that the tail of the KM curves is based upon a much smaller sample size than is available to the updated primary analysis: the number of patients at risk at 24 months is less than 40 in each treatment arm, compared to more than 100 at risk at 24 months in each arm in the updated primary analysis. These data are also problematic since they do not reflect what might have been observed, had different criteria been used prospectively to determine progression. Estimated mean time to progression is consequently upwardly biased.[2]

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

The MS examines the fit of various parametric survival models to its chosen PFS data. A gamma model provided the best fit to the treatment arm, while a log-logistic model provided the best fit to the comparator arm (MS section 7.3.1.1, Table 40, p.135). The MS notes that it is common to choose the same parametric form for both treatment arms but does not adopt a fully parametric form in the economic model; this uses a Kaplan Meier PFS until convergence of treatment and comparator arms, and a log-logistic model thereafter. This decision is justified based on 'visual inspection' and comparison with published PFS curves for patients with similar characteristics as described in du Bois and colleagues[23] (MS section 7.3.1.1, p.136). The ERG notes that the gamma model was the second best-fit parametric model to the comparator PFS data and may therefore have been a reasonable parametric model to use for both treatment arms, instead of the Kaplan Meier curve. The MS examines a gamma model in a deterministic sensitivity analysis for which results are presented (MS table 63, p.186) and discussed (MS section 7.7.10, p.192). The use of the convergence of the PFS curves to signal the switch from Kaplan Meier to parametric form at 28 months appears *ad hoc* and no justification is given. As noted above there are very low numbers in both treatment arms even before this time and the associated Kaplan Meier survival estimates may not be precise. Kaplan Meier PFS estimates were not examined in sensitivity analysis. The use of the log logistic model to extrapolate the tail of the Kaplan Meier curve, rather than a gamma model, is also not justified in the MS and not examined in sensitivity analysis.

The MS reports that in the GOG-0218 trial the secondary outcomes were OS, ORR, AEs and HRQoL (section 3). Three of these outcomes, OS, AEs and HRQoL, are used in the model. OS is considered here; AEs and HRQoL are considered in section 4.2.5 below.

OS occurs in the model for patients in the PFS and post-progression health states. Separate weekly probability of death estimates were calculated for the PFS state and progression states. Within a health state, the same weekly probability of death is applied to both treatment and comparator arms, although different probabilities are allowed for the post-progression health state in a sensitivity analysis (MS Table 63, p.186). The method for deriving weekly probability of death whilst in PFS is not stated in the MS. The ERG requested clarification of the methods and parameters used in the model for this calculation. The manufacturer clarified that the weekly probability of death whilst in PFS was derived from GOG-0218 trial data and estimates of all-cause mortality published by the Government Actuary's Department. In any cycle the probability of death from PFS is the maximum of these two mortality rates.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

The method for deriving a weekly probability of death estimate for patients in the disease progression health state is described (MS section 7.3.1.2 p.137). The MS states that these parameter estimates were derived from the probability of post-progression death, using data for patients' time from progression to death from the GOG-0218 trial, based on an assumption of constant probability of death (independent of time since progression). The original data from the GOG-0218 trial used in the model are not reported, so the validity of the resulting OS curve (which is shown in MS Figure 17, p.138) cannot be checked by the ERG.

In summary, the clinical effectiveness data used by the model have been obtained from a relevant trial and have an appropriate outcome. However the precise choice of data, functional forms and calculation of parameter values is in many cases not well justified and may not be robust.

4.2.5 Patient outcomes

The MS reports a systematic review of HRQoL studies (MS section 7.4.5, p.147). The search strategies (MS Appendix 10, p.238) are reported and appear adequate. Titles and abstracts were assessed using simple, appropriate selection criteria (MS Table 34, p125). However, it is not stated how many reviewers were involved. One potentially relevant study was identified and appraised in detail for HRQoL data (design not stated; appears to be a single-cohort study). The study reported EORTC utility data but the MS states that these were not appropriate for use in the cost effectiveness analysis as they had not been mapped to the EQ-5D.

The MS model incorporates HRQoL through the use of health state utility values for the PFS and disease progression health states. HRQoL was assessed in the GOG-0218 trial using FACT-O TOI and ADS scales and in the ICON7 trial using QLQ-C30, QLQ-OV28 and EQ-5D scales. The EQ-5D is recommended by NICE for assessments of utility.[24] The MS states (p.145) that EQ-5D scores from the ICON7 trial were therefore used in all model analyses.

A log-rank test confirmed that EQ-5D values did not differ between intervention and comparator arms for progression-free patients so an assumption was made that the utility values from both study arms at each time point could be combined (the method of data combination is not stated). Original EQ-5D scores are not provided in the clinical effectiveness section of the MS,

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

or in the primary publication for the ICON7 trial or its supplementary appendix[1] and so cannot be checked by the ERG. The MS does not mention whether there were any missing HRQoL data or how missing data were accounted for in analyses. The utility estimates for patients in PFS are shown in MS Table 46 (p.145).

The HRQoL utility value for disease progression was estimated from the mean utility in the ICON7 trial (0.7248) since EQ-5D scores were not routinely available for patients whose disease had progressed. The MS acknowledges that this data point is based on relatively few observations but justifies the estimate as being comparable to utility data from a trial of trabectedin which included refractory patients on second-line therapy with metastatic ovarian cancer.[17] The MS argues that the population in the trabectedin trial may be considered to have more severe disease than the whole of the population that relapse following bevacizumab in the ICON7 trial.

AEs in the model are assumed to be captured within the assessment of HRQoL (MS section 7.4.8, p.152). Specifically, the MS states that since the EQ-5D assessment of HRQoL was administered to patients at regular intervals before disease progression, it is expected that any impact on HRQoL by an AE has been captured and is reflected in the overall utility score. The ERG notes that the EQ-5D data used in the model are from the ICON7 trial, which employed a lower dose of bevacizumab than in the NICE scope. Any AEs caused by the higher dose of bevacizumab as specified in the NICE scope would not be captured using the utility data from the ICON7 trial.

4.2.6 Resource use

The resource use considered by the MS falls into three broad categories: drug acquisition; drug administration and pharmacy; and health state. A systematic search for resource data was conducted (MS section 7.5.3, p.155; Appendix 10.10). The databases searched and inclusion/exclusion criteria used are reported and appear appropriate. However, search strategies are not reported and the number of reviewers involved is not specified. No studies were found to be relevant.

The estimation of dosage and frequency of administration of bevacizumab was based on the SPC which is a dose of 15 mg per kilogram of body weight given once every 3 weeks.[25] The mean weight of a cohort of UK ovarian cancer patients (Sacco and colleagues[22]) was used to

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

calculate an expected dose per patient. Treatment duration was defined by observations from the GOG-0218 trial,[2] implemented as a Kaplan Meier survival curve of time on treatment (MS Figure 22, p.164). Dosing and frequency of comparator paclitaxel and carboplatin treatment were also based on their respective SPCs. Mean body surface area, age and weight measurements required to calculate an expected dose of paclitaxel and carboplatin per individual were again taken from the Sacco and colleagues retrospective cohort study.[22]

Dosing assumptions are consistent with the GOG-0218 trial. The MS states that the base case assumes no vial sharing of bevacizumab but vial sharing of carboplatin and paclitaxel as they are in more routine use (MS section 7.5.5.1, p.159). Overall the assumptions regarding patient size appear reasonable and do not have a great impact on the ICER. However the adopted dose is not consistent with current clinical practice in the UK which uses a bevacizumab dose of 7.5 mg/kg (ERG expert opinion).

Resource use estimates for the intervention and comparator are based on an assumed course of outpatient hospital treatment, including the pharmacy costs of drug preparation and administration (MS sections 7.5.5.5 – 7.5.5.6, p.162-163).

Treatment administration resource use estimates were based on treatment every three weeks as indicated by the SPC. In cycles where bevacizumab is administered together with carboplatin and paclitaxel additional pharmacy time of 12 minutes was considered to be the only extra resource, as determined by a prospective time and motion study for oxaliplatin[26] (MS section 7.5.5.6, p.162). Cycles where only bevacizumab is administered were assumed to incur a total pharmacy time of 12 minutes (MS section 7.5.5.5, p.162).

Clinical expert advice was used to inform PFS supportive care use (MS section 7.5.6, p.155) although the number of experts consulted is not made clear and no methods for pooling evidence are described. The Health Technology Assessment of Trabectedin[17] was used to provide estimates of weekly supportive care costs whilst in the PD state. Post-progression drug acquisition costs were not included in the model (MS section 7.5.8.1, p.169) as this information was not available in sufficient detail from the GOG-0218 trial.

Overall, the ERG considers that the relevant resource use appears to have been considered. The approach adopts the NHS perspective and is consistent with the reference case.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

4.2.7 Costs

All drug costs are current to 2012 and obtained from the British National Formulary[18] (bevacizumab) or DH Commercial Medicines Unit[27] (paclitaxel and carboplatin). On-treatment management and monitoring costs were taken from NHS reference costs for 2010/11[20] and PSSRU.[19] NHS reference costs for 2010/11 were also used for AEs (MS section 7.5.7, p.166). It was assumed that certain AEs do not involve significant cost to the health service and only those events occurring in greater than 2% of patients at grade 3 or 4 severity were incorporated into the analysis. All AEs were assumed to occur in the first cycle of the model, although an explanation is not provided in the MS. Although this is unlikely to reflect the clinical situation since bevacizumab would not be administered until the second cycle, it would not alter the overall costs included in the model.

The model uses a bevacizumab cost of £2,229 per patient per cycle. This is based on the average number of vials using UK patient attributes estimated by Sacco and colleagues[22] (MS Table 50, p60). Using average patient attributes from Sacco and colleagues[22] the cost of paclitaxel per patient per cycle used by the model is £21.80 and the cost of carboplatin is £18.51. Total model output costs by treatment arm for each clinical outcome are given in Table 12 below.

Table 12: Model cost outputs by clinical outcome (GOG-0218)

Comparator	Outcome	Cost (£)
Bevacizumab + chemotherapy	PFS	32,588
	PD	5,417
	Palliative care	6,248
	Total	44,254
Carboplatin + paclitaxel	PFS	5,281
	PD	5,593
	Palliative care	6,292
	Total	17,166

The ERG considers that all relevant costs appear to have been included and are transparently calculated. The manufacturer's approach is consistent with the NICE reference case.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

4.2.8 Consistency/ Model validation

Internal consistency

The economic model was developed in Microsoft Excel, with two alternative versions submitted for the analyses relating to the GOG-0218 and ICON7 RCTs. Random checking of the model has been done for some of the key equations in the model. The ERG has not undertaken a comprehensive check of all cells in the model. The model was checked to see if results were in the expected directions and had expected magnitude for changes to the model input parameters. The electronic model is fully executable, and inputs changed on the 'Model Inputs' worksheet produce changes in the deterministic results in the 'Results Table' worksheet. These can be used to replicate the results presented in the MS and the deterministic sensitivity analyses for the base case model, as reported in MS Table 63, p.186.

The model is generally well presented and user friendly, with most of the input parameters presented in the 'Model Inputs' worksheet. The ERG views the model as a reasonable approach to modelling the cost effectiveness of bevacizumab and from random checking the 'wiring' of the model appears to be accurate.

The MS provides a summary of the model results compared with clinical data in Table B4, p.180, shown here in Table 13. The MS does not provide any discussion on the differences between the clinical trial results and the model results. The ERG considers that the GOG-0218 model results are consistent with the clinical trial results for PFS. For OS, the ERG notes that there is a similar OS in both the chemotherapy and bevacizumab trial arms, whereas in the model the OS for bevacizumab is 2 months longer than for the chemotherapy arm.

Table 13: Summary of model results compared with clinical data for GOG-0218 trial

Outcome	Clinical trial result (median months)	Model result
Chemotherapy arm		
PFS	12.12	12.00
Post-progression survival	27.27	33.00
OS	39.39	45.00
Bevacizumab arm		
PFS	18.79	19.00
Post-progression survival	20.96	28.00
OS	39.75	47.00

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

The ERG notes that the treatment duration used within the model was a maximum of one year, rather than 15 months as stated in the SPC for bevacizumab for the GOG-0218 trial, and therefore the cost of bevacizumab has been underestimated. The total treatment cost of bevacizumab in the MS is £26,361 at a cost of £2229 per cycle, i.e. a mean treatment duration of 11.8 cycles, compared with the expected number of 13.7 cycles in the GOG-0218 trial (reported in MS Table 1, p.13). The ERG provides an analysis with longer treatment durations in section 4.3.

The ERG has uncovered the following coding errors in the MS model, which have a negligible effect on the model results. On sheet 'Bevacizumab + chem' and 'Chem', cells T2 to AG2 have calculated the total sum of their column incorrectly (should read =SUM(OFFSET(T2,**4**,,t_horizon*52,1)), instead of SUM(OFFSET(T2,**3**,,t_horizon*52,1))).

External consistency

The MS states that the results from the manufacturer's model are broadly consistent with the published literature found in their review of cost effectiveness studies, with the caveat that the published studies were for non-UK based healthcare systems. The ERG notes that the cost effectiveness results for GOG-0218 varied in the published cost effectiveness analyses from $326,500 (£200,000) per QALY gained[14] to $401,100 (£246,500) per progression-free life year saved.[13]

The OS estimates from the model were compared in the MS to estimates from an external source using ovarian cancer patients with similar disease severity and surgical outcome (MS Figure 17, p.138). The MS reports that the results from the model overestimate the survival of patients receiving chemotherapy after approximately 30 months.

4.2.9 Assessment of Uncertainty

One-way sensitivity analyses

The MS reports that a series of one-way sensitivity analyses were carried out on the base case model. The MS provided no rationale for the choice of variables included or excluded. The following variables were subjected to sensitivity analyses: PFS extrapolation, post-progression survival, utility values, administration, AE and chemotherapy costs, time horizon and discounting

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

rates. Some key input parameters (such as the cost of bevacizumab, treatment duration and variation in effectiveness) which might be expected to be highly influential on the cost effectiveness results have been omitted from the sensitivity analysis.

According to the MS, the key drivers of the cost effectiveness results are the dose and the cost of bevacizumab and the duration of treatment but these have not been shown in sensitivity analyses. From the values presented in MS Table 63/64, p.186/7, the model is also sensitive to the time horizon, and the distribution used for PFS. The model was insensitive to other parameters related to disease management costs for PFS and PD health states, AE costs, and health state utility values.

Scenario Analysis

The MS presents scenario analyses for vial sharing and trial patient characteristics (MS p.191) but no discussion or conclusion is given. In the base case analyses the MS does not use vial sharing. The MS states that in some centres vial sharing may be possible and this impacts on the expected cost per patient. The vial sharing scenario uses a reduced cost of bevacizumab of £2,109 per dose compared to £2,229 in the base case analysis. The result of this reduction in the cost of bevacizumab is a reduced ICER of £136,513 per QALY gained.

The trial patient characteristics scenario analysis uses the drug usage based upon the demographics from the RCTs. The mean body weight of women recruited to GOG-0218 is more than 10 kg more than the mean weight of UK ovarian cancer patients. The drug cost using the trial characteristics was £2,583 per dose compared to £2,229 in the base case analysis. The result of this increase in the cost of bevacizumab is an increased ICER of £166,287 per QALY gained.

Probabilistic Sensitivity Analysis

The PSA is run by clicking on the psa_run macro and takes about 5 minutes to run 5000 simulations. The results from the PSA are presented in the ‘Results Table’ worksheet. The results of the PSA (for GOG-0218) are shown as a scatterplot with the incremental cost and QALYs of bevacizumab shown (MS Figure 25, p.188). The MS presents PSA results for bevacizumab versus chemotherapy on MS p.189. These results are consistent with the deterministic base case results. The MS states that there is a 0% chance that the addition of 15

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

mg/kg bevacizumab to standard carboplatin and paclitaxel chemotherapy being considered cost effective at a willingness to pay threshold of £30,000 per QALY.

No explanation or rationale was given in the MS for the variables included in the PSA. Input variables and distributions in the PSA are shown in MS Table B2, p.142. The following input parameters were used in the PSA: utility values, costs and frequency of AEs, weekly supportive care costs in both PFS and progressed health states. According to the manufacturer's letter of clarification, parameter estimates for the parametric PFS and OS functions were also included in the PFS. However, the Kaplan Meier survival estimates of PFS and OS taken directly from the clinical trials were not subject to uncertainty analysis.

The ERG notes the lack of variability in the simulations results. The ERG considers that the PSA does not include all the uncertainty of the model and that key parameters have been omitted from the PSA, for example the cost of bevacizumab and the clinical effectiveness of bevacizumab.

4.2.10 Comment on validity of results with reference to methodology used

The manufacturer submitted two electronic models, for the GOG-0218 and ICON7 RCTs. The ICON7 trial reflects a lower dose of bevacizumab than the NICE scope (7.5 mg/kg every 3 weeks).

The structure adopted for the economic model is reasonable, and consistent with previous economic evaluations developed for advanced cancer. The methods of analysis are generally appropriate and conform to NICE methodological guidelines.

The parameters used for the model are generally appropriate. The population used in the model is that from the relevant trial (GOG-0218) and is generally representative of those treated in secondary care in the UK, although the population may not fully represent patients who have had comorbidities. An error was detected for the treatment duration of bevacizumab used in the model, which underestimated the total treatment costs. These have been documented in this report, along with corrected results in section 4.3.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

4.3 Additional work undertaken by the ERG

The ERG has run additional analyses using the manufacturer's GOG-0218 model for changes to treatment duration, treatment cost and time horizon (Table 14).

In section 4.2.8, the ERG noted that the treatment duration used within the model was a maximum of one year, rather than 15 months as stated in the SPC for bevacizumab for the GOG-0218 trial. Using the trial discontinuation rates in the GOG-0218 trial and with treatment for a maximum of 15 months the ICER of bevacizumab increased from the base case of £144,066 per QALY gained to £160,788 per QALY gained.

The ERG investigated the effect of changing the treatment cost of bevacizumab to the treatment cost for the lower dosage of 7.5 mg/kg, commonly used in clinical practice, assuming the same treatment effect as seen in the GOG-0218 trial. For the lower bevacizumab dosage, the ICER of bevacizumab reduced to £77,884 per QALY gained.

The ERG noted in section of 4.2.1 of this report that the MS model adopted a time horizon of ten years. The ERG investigated the effect of changing the time horizon to the maximum permitted in the MS model of 25 years. For this analysis, the ICER reduced from the base case of £144,066 per QALY gained to £127,701 per QALY gained.

Finally the ERG combined the analyses above for treatment duration of 15 months and a time horizon of 25 years which produced an ICER similar to the base case of £142,477 per QALY gained.

Table 14: ERG analyses on effect of changes to treatment duration, treatment cost and time horizon on model results

Scenario	Treatment	Mean total costs, £	Mean QALYs	ICER (£/QALY gained)
Base case	CPB	44,254	3.16	-
	CP	17,166	2.97	-
	Incremental	27,089	0.19	144,066
Total treatment duration 15 months using trial discontinuation rates	CPB	47,399	3.16	-
	CP	17,166	2.97	-
	Incremental	30,233	0.19	160,788
Reduced treatment cost using cost of 7.5mg/kg (£1177)	CPB	31,810	3.16	-
	CP	17,166	2.97	-

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

	Incremental	14,645	0.19	77,884
Time horizon of 25 years	CPB	45,174	3.342	-
	CP	18,001	3.129	-
	Incremental	27,173	0.21	127,701
Treatment duration 15 months, and time horizon of 25 years	CPB	48,318	3.342	-
	CP	18,001	3.129	-
	Incremental	30,317	0.21	142,477

4.4 Summary of uncertainties and issues

- The treatment duration used within the model has been underestimated by using a maximum of one year, rather than 15 months as stated in the SPC for bevacizumab and for the GOG-0218 trial, and therefore the cost of bevacizumab has been underestimated.
- The economic model has not used a lifetime time horizon and has instead used a time horizon of 10 years, which may not be long enough to reflect all differences in costs or outcomes between the treatments.
- The MS has not included all model parameters in either the deterministic or probabilistic sensitivity analyses and so the full uncertainty around the model results has not been shown. In particular key parameters associated with the cost and effectiveness of bevacizumab have been omitted.
- A range of hazard ratios for PFS have been presented by the manufacturer using different methods of assessment. Of these, a relatively favourable hazard ratio for PFS was used in the model, which might have produced a more favourable cost effectiveness estimate.

5 End of life

NICE end of life treatment criteria were not applicable and not included in the MS.

6 Innovation

The manufacturer's case for innovation states that bevacizumab is the first licensed anti-VEGF targeted therapy in ovarian cancer. The induction and growth of ovarian cancers have been

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

linked with high levels of VEGF and the expression of VEGF receptors on ovarian cancer cells. Bevacizumab directly targets VEGF-driven angiogenesis to reduce vascularisation of tumour and thereby inhibit tumour growth. Its AE profile, unlike that of cytotoxic agents, allows it to be combined with cytotoxic chemotherapies without providing an intolerable additional burden of toxicity. This directly targeted therapy plus different toxicity profile represents an innovative step change in the management of ovarian cancer.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The MS includes evidence on the clinical effectiveness of bevacizumab in combination with carboplatin and paclitaxel for the treatment of advanced ovarian cancer from two RCTs. Results presented in the MS suggest that first-line therapy with bevacizumab in combination with chemotherapy followed by bevacizumab maintenance is superior for PFS than chemotherapy plus placebo and appear to be unbiased estimates of effectiveness. A range of estimates is presented so the exact effect size is not clear. The licensed dose used in the key trial is not the same as that used in current clinical practice.

7.2 Summary of cost effectiveness issues

The MS includes evidence on the cost effectiveness of bevacizumab and chemotherapy compared to chemotherapy alone for advanced ovarian cancer. The model structure and methods adopted for the economic evaluation are reasonable and are generally appropriate. The model structure and model parameter inputs are consistent with the clinical disease pathways and the available clinical trial evidence. However, it should be noted that a relatively favourable hazard ratio for PFS was used in the model (based on an analysis that censored progression events defined by rising CA-125). The model results suggest that bevacizumab is not cost effective at the licensed dose for a willingness-to-pay threshold of £20,000-£30,000 per QALY.

The MS has not included all model parameters in either the univariate or probabilistic sensitivity analyses and so the full uncertainty around the model results has not been shown. The results produced for the licensed dose underestimate the treatment duration and cost for bevacizumab.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

8 REFERENCES

(1) Perren TJ, Swart AM, Pfisterer J, Ledermann JA, Pujade-Lauraine E, Kristensen G et al. A phase 3 trial of bevacizumab in ovarian cancer. *N Engl J Med* 2011; 365(26):2484-2496.

(2) Burger RA, Brady MF, Bookman MA, Fleming GF, Monk BJ, Huang H et al. Incorporation of bevacizumab in the primary treatment of ovarian cancer. *New England Journal of Medicine* 2011; 365:2473-2483.

(3) Phase III trial of bevacizumab (BEV) in the primary treatment of advanced epithelial ovarian cancer (EOC), primary peritoneal cancer (PPC), or fallopian tube cancer (FTC): A Gynecologic Oncology Group study. 2010.

(4) Burger RA, Brady MF, Bookman MA, Monk BJ, Walker JL, Homesley HD et al. Safety and subgroup efficacy analyses in GOG-0218, a phase III trial of bevacizumab in the primary treatment of advanced epithelial ovarian cancer, primary cancer, or fallopian tube cancer: A Gynecologic Oncology Group study. ESMO abstract [978PD]. 2010.

(5) Monk BJ, Huang H, Burger RA, Mannel RL, Homesley HD, Fowler J et al. Quality of life outcomes of randomized, placebo controlled trial of beacizumab in the front line treatment of ovarian cancer: A Gynecologic Oncology Group Study. ECCO/ESMO [23LA]. 2011.

(6) Perren T, Swart A, Pfisterer J, Ledermann J, Lortholary A, Kristensen G et al. ICON7: A phase III randomised gynaecologic cancer intergroup trial of concurrent bevacizumab and chemotherapy followed up by maintenance bevacizumab, versus chemotherapy alone in women with newly diagnosed epithelial overian (eoc), primary peritoneal (ppc) or fallopian tube cancer (ftc). *ESMO* 2010;(LBA4).

(7) Result of interim analysis of overall survival in the GCiG ICON7 phase III randomized trial of bevacizumab in women with newly diagnosed ovarian cancer. 2011.

(8) Stark D, Nakivell M, Hipert F, Elit L, Brown J, Lanceley A et al. Quality of life data in the ICON7 GCIC phase III randomised clinical trial. ECCO/ESMO [22LBA]. 2011.

(9) Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical Oncology* 1993; 11(3):570-579.

(10) Basen-Enqquist K, Bodurka-Bevers D, Fitzgerald MA, Webster K, Cella D, Hu S et al. Reliability and validity of the functional assessment of cancer therapy - ovarian. *Journal of Clinical Oncology* 2001; 19(6):1809-1817.

(11) Wenzel L, Huang HQ, Cella D, Walker JL, Armstrong DK, Gynecologic Oncology Group. Validation of the FACT/GOG-AD subscale for ovarian cancer-related abdominal discomfort: A Gynecologic Oncology Group study. *Gynecologic Oncology* 2008; 110(1):60-64.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

(12) Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. Third edition. 2009. York Publishing Services Ltd., CRD.

(13) Cohn DE, Kim KH, Resnick KE, O'Malley DM, Straughn JM, Jr. At what cost does a potential survival advantage of bevacizumab make sense for the primary treatment of ovarian cancer? A cost-effectiveness analysis. *J Clin Oncol* 2011; 29(10):1247-1251.

(14) Lesnock JL, Farris C, Krivak TC, Smith KJ, Markman M. Consolidation paclitaxel is more cost-effective than bevacizumab following upfront treatment of advanced epithelial ovarian cancer. *Gynecol Oncol* 2011; 122(3):473-478.

(15) Economic evaluation of bevacizumab for the treatment of advanced ovarian cancer in Mexico. 2012.

(16) Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; 313(7052):275-283.

(17) Papaioannou D, Rafia R, Stevenson MD, Stevens JW, Evans P. Trabectedin for the treatment of relapsed ovarian cancer. *Health Technol Assess* 2011; 15 Suppl 1:69-75.

(18) Joint Formulary Committee. British National Formulary 63. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 2012.

(19) Curtis L. Unit Costs of Health and Social Care 2011. http://www pssru ac uk/pdf/uc/uc2010/uc2010 pdf [2011

(20) Department of Health. NHS reference costs 2010-2011. http://www dh gov uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_122803 [2011 [cited 2010 June 7];

(21) Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes (3rd edition). *Oxford University Press* 2005.

(22) Sacco JJ, Botten J, Macbeth F, Bagust A, Clark P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. *PLoS One* 2010; 5(1):e8933.

(23) du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). *Cancer* 2009; 115(6):1234-1244.

(24) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. *NICE, London* 2008.

(25) Roche Products Ltd. Avastin Summary of Product Characteristics. http://www medicines org

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

uk/EMC/medicine/15748/SPC/Avastin+25mg+ml+concentrate+for+solution+for+infusion/ [2012

(26) PCN74 A Service Evaluation to compare secondary care resource use between XELOX and FOLFOX-6 regimens in the treatment of metastatic colorectal cancer (MCRC) from a UK National Health Service (NHS) perspective. 2008.

(27) Department of Health. Commercial medicines unit. Electronic market information tool [2012 Available from: URL:http://cmu.dh.gov.uk/electronic-market-information-tool-emit/

9 APPENDICES

9.1 Appendix 1: ICON 7

Description and context of ICON7 study

The ICON7 trial[1] is a two armed, open-label, phase III RCT which compared the effectiveness of bevacizumab at a dose of 7.5mg/kg combined with carboplatin (AUC5 or 6) and paclitaxel (175mg/m^2) (q3w for 6 cycles) followed by extended bevacizumab for 12 cycles (the 'CPB7.5+' arm) with the effectiveness of chemotherapy with carboplatin (AUC5 or 6) and Paclitaxel (175mg/m^2) alone (q3w for 6 cycles) (the 'CP' arm) in the treatment of ovarian cancer. The study included patients with both early and advanced stage disease, and 81% of the patients entered into the trial had advanced ovarian cancer (FIGO stage III and stage IV disease). As the ICON7 study included patients with early stage ovarian cancer and used a bevacizumab dose of 7.5 mg/kg, the ERG noted that it did not match the scope or the licensed indication and did not meet the patient population specified in the inclusion criteria for the review in the MS (patients with advanced ovarian cancer). Based on this, we restricted our review of the evidence submitted by the manufacturer to the GOG-0218 trial in the main part of this report. We have, however, added commentary on the quality of the ICON7 trial and its key findings, including an exploratory subgroup analysis of PFS in patients who were Stage III sub-optimally debulked and Stage IV debulked (who most closely resemble the patient population in the GOG-0218 trial used for the licensing authorisation), here for information.

Manufacturer and ERG assessment of trial quality

The MS provides a quality assessment of the ICON7 trial in Appendix 3 (MS p.218) and a summary assessment is tabulated in Section 6.4 of the MS (p.78). The quality assessment

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

follows the NICE criteria and is appropriate. Table 15 shows the ERG independent assessment of study quality and the MS assessment. As this table shows, the ERG assessment partly agrees with that of the manufacturer.

Table 15: Manufacturer and ERG assessment of trial quality of ICON7 study

NICE QA Criteria for RCT	MS response	ERG response
1. Was the method used to generate random allocations adequate?	Yes	Yes
2. Was the allocation adequately concealed?	No	Yes. MS states 'no' as the trial was open-label. However, this QA question refers to whether or not group allocation could have been foreseen prior to randomisation, and the ERG notes that allocation was adequately concealed.
3. Were the groups similar at the outset of the study in terms of prognostic factors, e.g. severity of disease?	Yes	Yes
4. 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)?	No	No
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Stated as 'no' on MS p. 78 and as 'not clear' on MS p. 218.	Yes. Proportionally more patients in the CPB7.5+ arm than in the CP arm were withdrawn from treatment (26.2% and 9.8%, respectively, MS p76). Reasons are not provided for all patient withdrawals in the MS, but are provided in the trial paper.[1]. The MS states that the proportion of patients withdrawn due to insufficient therapeutic response or death was higher in the CPB7.5+ arm than in the CP arm (12.8% of patients in the CPB7.5+ arm and 2.4% of patients in the CP arm; one patient in the CPB7.5+ arm died and two patients in the CP arm died). The trial paper[1] shows that more patients in the CPB7.5+ arm than in the CP arm withdrew due to an AE or intercurrent illness. This was adjusted for by use of ITT analyses of the PFS and OS outcomes.
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
7. Did the analysis include an	Yes	Yes. ITT analysis of the PFS and OS

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

intention to treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?		outcomes were conducted.

Results

Primary outcome PFS

Results for the primary outcome of PFS are shown in Table 16. For the ITT population the risk of disease progression or death was decreased by 13% for patients in the CPB7.5+ arm compared with the CP arm (HR 0.87; CI 0.77, 0.99, p=0.04).

Exploratory subgroup analysis of PFS in patients who were Stage III sub-optimally debulked and Stage IV debulked is presented as these groups of patients most closely resemble those in the GOG-0218 trial used for the licensing authorisation. There was a statistically significant increase in median PFS of 5.5 months in the CPB7.5+ arm compared to the CP arm (CP 10.5 months, CPB7.5+ 16.0 months; HR 0.73, 95%CI 0.0.60, 0.93, p=0.002). (See MS Table 18, p.91).

Subgroup analysis by disease stage and debulking status are also presented (MS Table 20, p.93) which show a statistically significant increase in PFS for Stage III sub-optimally debulked disease with CPB7.5+ compared with CP. No significant improvement was shown for the other groups presented. A second subgroup analysis by disease stage and outcome of surgery shows that for patients with Stage III inoperable and Stage IV disease the hazard ratio for PFS was 0.66 (95%CI 0.48, 0.91, p=0.011). (MS Table 19, p.92)

Table 16: Median PFS for overall study results and exploratory subgroup analyses (investigator assessed, updated analysis) from Study ICON7

		CP (n=764)	**CPB7.5+ (n=764)**
Investigator assessment (ITT population)	Median PFS mths	17.4	19.8
	Hazard ratio (95% CI)		0.87 (0.77, 0.99)
	P value		0.04
		CP (n=234)	**CPB7.5+ (n=231)**
Stage III sub-optimal and Stage IV debulking ('high risk', similar to GOG-0218)	Median PFS mths	10.5	16.0
	Hazard ratio (95% CI)		0.73 (0.60, 0.93)
	P value		0.002
Stage III optimal debulking*		**CP (n=368)**	**CPB7.5+ (n=383)**
	Median PFS mths	17.7	19.3
	Hazard ratio		0.89

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

	(95% CI)		(0.74, 1.07)
	P value		n/a
Stage III sub-optimal debulking		**CP (n=154)**	**CPB7.5+ (n=140)**
	Median PFS mths	10.1	16.9
	Hazard ratio (95% CI)		0.67 (0.52, 0.87)
	P value		n/a
Stage IV		**CP (n=97)**	**CPB7.5+ (n=104)**
	Median PFS mths	10.1	13.5
	Hazard ratio (95% CI)		0.74 (0.55, 1.01)
	P value		n/a
Inoperable Stage III and Stage IV		**CP (n=106)**	**CPB7.5+ (n=106)**
	Median PFS mths	Not provided	Not provided
	Hazard ratio (95% CI)		0.66 (0.48, 0.91)
	P value		0.011

*with or without gross residual disease
Clarification requested from the manufacturer stated that missing p values are not available (n/a)

Secondary outcomes

Overall survival

Early ITT analysis of OS showed in favour of the CPB7.5+ arm although median duration of OS could not be determined as data were not mature enough at the time of data cut-off.

Interim OS from the ICON7 'high risk' subgroup analysis shows that Stage III sub-optimally debulked and Stage IV patients had a median OS improvement of 7.8 months in the CPB7.5 arm compared with the CP arm which was statistically significant (CP 28.8 months, CPB7.5+ 36.6 months; HR 0.64, p=0.002; MS Figure 12, p.95). The HR indicates a 36% reduction in risk of death in the high risk patients treated with CPB7.5+ compared with CP patients. OS data are shown in Table 17.

Table 17: Overall survival for ITT population and exploratory 'high risk' subgroup (ICON7)

		CP (n=764)	**CPB7.5+ (n=764)**
Investigator assessed (early ITT population)	Median mths	Not reached	
	Hazard ratio (95% CI)		0.85 (0.69, 1.04)
	P value		0.1167
		CP (n=234)	**CPB7.5+ (n=231)**
ICON7 Stage III sub-optimal debulking and Stage IV	Median mths	28.8	36.6
	Hazard ratio (95% CI)		0.64 (0.48, 0.85)
	P value		n/a

Clarification requested from the manufacturer stated that missing p value not available (n/a)

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Objective response rate

Investigator assessed ORR from the ICON7 study was statistically significantly higher in the CPB7.5+ arm compared with the CP arm; however, no data are presented for the relevant subgroups.

HRQoL

It is reported that some women receiving bevacizumab has a statistically significant but clinically small detriment in global QoL but no HRQoL data are presented for the relevant subgroups for the ICON7.

Adverse events

The safety analyses were not ITT analyses; patients were only included in the safety analyses if they received treatment. In the MS it is stated that patients were not permitted to cross-over in this study (MS p.38), yet the safety analysis section indicates that they did (MS p.69). In the safety analyses, patients treated with bevacizumab in the CP arm (including those treated with it due to error) were included in the CPB7.5+ arm for analysis, and patients not treated with bevacizumab in the CPB7.5+ arm were included in the CP arm for analysis. The proportion of patients who crossed over from each arm is not detailed in the MS. Although this approach breaks randomisation, it is unlikely to have affected the conclusions made about AEs associated with bevacizumab.

The MS provides a summary of the overall incidence of AEs for each arm (MS Table 31, p.109), including AEs leading to treatment discontinuation and death. This shows that proportionally more patients in the CPB7.5+ arm than in the CP arm experienced a grade 3-5 AE (64.6% and 54.3%, respectively), serious AE (37.7% and 23.5%, respectively) and AEs leading to treatment discontinuation (22.0% and 8.9%, respectively). A greater proportion of patients in the CPB7.5+ arm than in the CP arm also experienced an AE of special interest to bevacizumab (Table 18). Death rates were similar between the two arms. The MS also provides a summary of grade 3-5 AEs of special interest to bevacizumab in the overall safety population and in the subgroup of patients with Stage III sub-optimally debulked or Stage IV disease (who most closely matched the patient population in the GOG-0218 trial) in Table 33 (MS p.111). This shows that AE rates were similar between arms for both the overall population and advanced disease subgroup, except that:

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

- proportionally more patients in the overall population and the advanced disease subgroup analyses experienced hypertension in the CPB7.5+ arm than in the CP arm. (Overall population: 6.0% in CPB7.5+ arm and 0.3% in the CP arm. Advanced disease subgroup: 7.8% in CPB7.5+ arm and 0.4% in CP arm.)
- in the advanced disease subgroup analysis, proportionally more patients in the CPB7.5+ arm experienced wound healing complications than in the CP arm (6.7% and 0.4%, respectively).

Table 18: Number and proportion of patients in the CP and CPB7.5+ arms who experienced an AE of special interest to bevacizumab

AE	**CP** **N = 763**	**CPB7.5** **N = 746**
AE of special interest to bevacizumab	362 (47.4%)	552 (74.0%)
Grade 3-5 AE of special interest	156 (20.4%)	240 (32.2%)
Serious AE of special interest	49 (6.4%)	123 (16.5%)

Summary of clinical effectiveness

Although the PFS subgroup analyses generally concur with the overall study results, the specific effect sizes vary and should be viewed with caution as, although preplanned, they may not be powered to detect differences between treatment groups. Results suggest an OS benefit in high risk patients with 36% reduction in relative risk of death in patients treated with CPB7.5+ compared with CP patient (patients not allowed to cross-over post-progression).

Data from the ICON7 study support the results from Study GOG-0218 for PFS both for the ITT population and for a relevant subgroup of patients. However, the 'high risk' group does not completely match the patient group in the GOG-0218 study (it only covers 2 of the patient groups – Stage III sub-optimal debulking and Stage IV) as the definition of optimal debulking differs between the two studies.

Economic evaluation

The ICON7 model is a three state survival model which is built around the principal outcomes of the ICON7 study, PFS and OS (MS section 7.2.2.2, p.132). The ICON7 trial did not use bevacizumab at the licensed dose of 15mg/kg but rather at a dose of 7.5 mg/kg, and for a different treatment duration and in a different study population to the GOG-0218 trial. A different classification of PFS was also used. For these reasons it was felt inappropriate to combine the

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

results of the two trials (MS section 6.5.4, p.99) and consequently two separate economic models were built. As with the GOG-0218 economic model, the ICON7 model uses Kaplan Meier survival curves for PFS until the convergence of curves for the intervention and comparator. This occurs at month 24 in contrast to month 28 in the GOG-0218 trial (MS section 7.3.1.1, p.136). To extrapolate PFS survival times beyond clinical follow up, the model utilises a parametric survival model beyond month 24. Parametric functions are also used to describe long-term OS on the two treatment arms.

The results of the ICON7 trial include various Kaplan Meier curves for PFS and the MS notes that the model was developed using updated data from 30 November 2010 (MS section 7.2.2.2, p.132) for an 'expanded high risk cohort' (MS section 7.2.1.2, p.130). This cohort was chosen to reflect the licensed indication for bevacizumab. The MS notes that some patients in this cohort are outside the license but they represent less than 1% of the subgroup and so are unlikely to have a significant effect on clinical outcomes. The cohort however is comprised of only 495 patients at outset (MS Table 38, p.130), leading to small sample sizes in the tails of the KM curves, particularly for PFS, and therefore low precision in the survival estimates obtained for these times. At 24 months following randomization there are fewer than 40 patients remaining in each treatment arm for PFS.[6]

The MS examines the fit of various parametric survival models to its chosen ICON7 PFS data and presents results (MS Table 41, p.137). None of the models was felt to provide a satisfactory fit to the data (MS section 7.3.1.1, p.137) and so a Kaplan Meier PFS was used until convergence of treatment and comparator arms, and a log-logistic model thereafter. The precise grounds for rejecting the parametric model fits are not specified in the MS but fully parametric log-logistic and gamma models are examined in sensitivity analysis (MS Table 64, p.187) and neither is found to result in an appreciably worse ICER. As with the GOG-0218 model, the timing of the switch from KM PFS to parametric model, at convergence of KM curves, is not justified in the MS and is not examined in sensitivity analysis. The choice of a log-logistic model to provide the parametric tail to the KM PFS curves is also not justified and not examined in sensitivity analysis. MS Table 41 (p.137) reports that the log-logistic model is the best-fit to the PFS data but does not specify which treatment arm this relates to; it is unclear if the log-logistic model provided the best fit to both treatment arms. KM PFS estimates were not examined in sensitivity analysis.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

In contrast to the GOG-0218 model which uses the same weekly probability of death for both treatment arms, because of confounding, the ICON7 model uses different OS curves for each treatment arm. A gamma survival model was found to provide the best fit to the data (MS Table 44, p.139) but is not adopted as the economic model base case as, on the basis of discussions with ovarian cancer clinicians, the tail of this fit was felt to underestimate long-term survival (MS section 7.3.1.2, p.139). Instead the economic model adopts log-logistic fits to OS on the basis of comparisons with data from du Bois and colleagues.[23] The ERG has examined these data and finds that they represent an appropriate clinical subgroup, albeit one which has not followed a course of bevacizumab treatment. Uncertainty in the parameters of the log-logistic model is considered in probabilistic sensitivity analysis (MS section 7.6.3, p.178). Gamma and Weibull fits to the ICON7 OS data are examined in one-way sensitivity analysis reported (MS Table 64, p.187) the relevant extract of which is given in Table 19 below. It may be seen from the table that use of a gamma model for OS, rather than a log-logistic model, has a considerable impact on the ICER.

Table 19: Deterministic sensitivity analysis (ICON7)

Parameter	Base case value	Alternative	Incremental costs	Incremental QALYs	ICER
BASE CASE			£17,729	0.561	£31,592
OS	Log Logistic	Gamma	£17,667	0.475	£37,173
		Weibull	£17,846	0.539	£33,085

The MS states that the ICON7 model includes post-progression costs for selected treatments, in contrast with the GOG-0218 model which does not (MS section 7.5.8.1, p.169). Dosing assumptions for these treatments are taken from the appropriate SPC [25] whilst costs are taken from BNF 63.[18] The total assumed cost of post-progression treatment is given (MS Table 58, p.171). The ERG has examined these costs and feels that they have been robustly and transparently calculated, but notes that they are not currently incorporated into the model cost calculations. Total model costs by treatment arm for each clinical outcome are given in Table 20 below.

Table 20: Model output costs by treatment arm and clinical outcome (ICON7)

Comparator	Outcome	Cost (£)
Bevacizumab + chemotherapy	PFS	19,447
	PD	8,208
	Palliative care	6,190
	OS	33,846

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.

Carboplatin + paclitaxel	PFS	1,793
	PD	7,917
	Palliative care	6,406
	OS	16,116

When compared with Table 12 (section 4.2.7), Table 20 shows that apart from the additional cost in the bevacizumab treatment arms, the two trials differ in their costs in two main respects: there is a noticeably higher cost associated with PFS on the chemotherapy arm in the GOG-0218 trial (£5,281) than in the ICON7 trial (£1,793); and progression costs (PD) are higher on both treatment arms in the ICON7 trial.

The first of these differences may be attributed to the higher cost of AEs in the GOG-0218 chemotherapy arm. These amounted to £3,512, compared to £233 for the ICON7 trial, and the ERG notes that reporting of AEs may not have been consistent across the two trials (MS Table 26, Table 32).

The higher PD-state costs for the ICON7 trial compared to the GOG-0218 trial (Table 12 in section 4.2.7, Table 20) may be attributed to a greater proportion of patients being in the progression state, on average, in the ICON7 trial than in the GOG-0218 trial.

Overall, the ICON7 model is built using appropriate data, with appropriate outcomes and is generally well described and justified in the MS. The data used in the model do however represent a subgroup of a clinical trial and consequently estimates based on this subgroup may not be very precise, simply because of the relatively small sample size. The ICER appears particularly sensitive to assumptions of parametric form for OS.

Copyright 2012 Queen's Printer and Controller of HMSO. All rights reserved.