BEVACIZUMAB IN COMBINATION WITH FLUOROPYRIMIDINE-BASED CHEMOTHERAPY FOR THE FIRST-LINE TREATMENT OF METASTATIC COLORECTAL CANCER – A SINGLE TECHNOLOGY APPRAISAL

Produced by School of Health and Related Research (ScHARR), The University of Sheffield.

Authors Sophie Whyte, Research Associate, ScHARR, University of Sheffield, [redacted]

Abdullah Pandor, Research Fellow, ScHARR, University of Sheffield, [redacted].

Matt Stevenson, Senior Operational Researcher, ScHARR, University of Sheffield, .[redacted]

Angie Rees, Information Specialist, ScHARR, University of Sheffield, [redacted]

Correspondence to Sophie Whyte, Research Associate, ScHARR, University of Sheffield, [redacted]

Date completed 25th September 2009

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/211/01.

Declared competing interests of the authors: None.

Acknowledgements

Dr Mark Saunders, [redacted]

Dr Debra Furniss, [redacted]

We would also like to thank Gill Rooney, Project Administrator[redacted], for her help in preparing and formatting the 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.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

List of abbreviations

APAS	Avastin Patient Access Scheme
CCTR	Cochrane Controlled Trials Register
CEAC	Cost-Effectiveness Acceptability Curve
CI	Confidence Interval
CIC	Commercial-In-Confidence
CRC	Colorectal cancer
CVAD	Central Venous Access Device
EQ-5D	EuroQol-5D
ERG	Evidence Review Group
HRQoL	Health-Related Quality of Life
FOLFIRI	intravenous 5- fluorouracil plus folinic acid plus irinotecan
FOLFOX	intravenous 5- fluorouracil plus folinic acid plus oxaliplatin
5-FU	5-fluorouracil
HUI	Health Utility Index
ICER	Incremental Cost Effectiveness Ratio
ITT	Intention-To-Treat
mg/d	milligrams per day
mg/kg	Milligram per kilogram
MLE	Maximum Likelihood Estimate
MS	Manufacturer's Submission
NICE	National Institute for Health and Clinical Excellence
PD	Progressive disease
PFS_{PT}	Progression free survival post treatment
PFS_{T}	Progression free survival on treatment
PSA	Probabilistic Sensitivity Analyses
QALY	Quality Adjusted Life Year
RCT	Randomised Controlled Trial
RDI	Relative dose intensity
RECIST	Solid evaluation criteria in solid tumours
RR	Relative Risk
SEM	Standard Error of mean
STA	Single Technology Appraisal
VEGF	vascular endothelial growth factor
XELIRI	oral capecitabine plus intravenous irinotecan

XELOX	oral capecitabine plus intravenous oxaliplatin

1. SUMMARY

1.1 Scope of the submission

The manufacturer's submission (MS) generally reflects the scope of the appraisal issued by NICE, and is appropriate to the NHS. Although the majority of the MS reflects the use of bevacizumab in combination with oxaliplatin-based chemotherapy as first line therapy (for patients not previously treated for metastatic disease) in individuals with histologically confirmed metastatic colorectal cancer, it does not reflect the broader population outlined in the licensed indication and final scope issued by NICE. The licensed indication permits the use of bevacizumab in combination with fluoropyrimidine-based chemotherapy for the treatment of patients with metastatic colorectal cancer but does not specify a line of treatment. Whilst the NICE scope broadly reflects the licensed indication; the manufacturer is seeking approval for first line use only. The MS defines the intervention as bevacizumab in combination with oxaliplatin and either 5-fluoruracil or capecitabine for people with metastatic colorectal cancer for whom oxaliplatin-including chemotherapy regimens are suitable. The MS considered oxaliplatin-including chemotherapy regimens without bevacizumab as the most relevant comparator, as reflected in the scope. However, a comparison with irinotecan-including chemotherapy regimens without bevacizumab was considered of limited clinical and economic relevance because it is not commonly used within the UK. While the systematic review undertaken in the MS did not consider irinotecan-based chemotherapy regimens as a relevant comparator, the manufacturer undertook an economic evaluation with irinotecan-based chemotherapy (given the small number of patients for whom this comparison is relevant) for completeness (using data from a published, peer reviewed, mixed treatment comparison). The outcome measures identified in the scope were all relevant and included overall survival, progression free survival, response rate, adverse effects and health related quality of life (HRQoL). The results provided are presented in terms of cost per quality adjusted life year (QALY) with a time horizon of eight years, which is equivalent to a life time horizon in the population of interest, with the perspective of costs taken from a NHS and Personal Social Services perspective.

1.2 Summary of submitted clinical effectiveness evidence

- The manufacturer's submission to NICE includes a systematic review of the clinical effectiveness literature. Although two randomised controlled trials were identified, one as first line therapy (the NO16966 trial) and one in second-line therapy (for previously treated patients with metastatic disease, E3200 trial), the manufacturer is seeking approval for first line use only (and therefore the N016966 trial forms the main pivotal evidence in the submission). The manufacturer claims that they could not demonstrate a cost-effectiveness case for the use of bevacizumab in second-line therapy.

- The NO16966 trial was a phase III, multicentre, multinational, two-arm, randomised, open label study with the primary objective of confirming the non-inferiority of XELOX (oxaliplatin plus capecitabine) compared with FOLFOX-4 (oxaliplatin plus 5-fluorouracil and folinic acid) in adult patients with histologically confirmed metastatic colorectal cancer who had not previously been treated. Following randomisation of 634 patients, the open label study was amended to include a 2x2 factorial randomised (partially blinded for bevacizumab) phase III trial (n=1401) with the co-primary objective of demonstrating superiority of bevacizumab in combination with chemotherapy (B-XELOX or B-FOLFOX-4) compared with placebo (P-XELOX or P-FOLFOX-4). The dose of bevacizumab was 5 mg/kg every two weeks (B-FOLFOX-4) or 7.5 mg/kg every three weeks (B-XELOX).

- The manufacturers' primary pooled analysis of superiority (using the intention to treat population) in the NO16966 trial showed that after a median follow up of 28 months, the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly improved progression free survival and overall survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined) in adult patients with histologically confirmed metastatic colorectal cancer not previously treated (median progression free survival: 9.4 versus 7.7 months [absolute difference, 1.7 months]; hazard ratio, 0.79; 97.5% CI: 0.72 to 0.87; p=0.0001; median overall survival: 21.2 versus 18.9 months [absolute difference, 2.3 months]; hazard ratio, 0.83; 97.5% CI: 0.74 to 0.93; p=0.0019).

- A secondary pooled analysis of superiority, restricted to patients in the second 2x2 factorial part of the NO16966 study (as per the original statistical trial plan [B-XELOX / B-FOLFOX-4 combined versus P-XELOX/ P-FOLFOX-4 combined] and which the ERG believe to be more appropriate) found similar results (median progression free survival: 9.4 versus 8.0 months [absolute difference, 1.4 months]; hazard ratio, 0.83; 97.5% CI: 0.72 to 0.95; p=0.0023; median overall survival: 21.3 versus 19.9 months [absolute difference, 1.4 months]; hazard ratio, 0.89; 97.5% CI: 0.76 to 1.03; p=0.0769).

- The manufacturers' pooled analysis of non-inferiority (using the eligible patient population and the intention to treat population) showed that the XELOX (XELOX/P-XELOX/B-XELOX combined) and FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/ B-FOLFOX-4 combined) based regimens were equivalent for both progression free survival (p=not significant, values not reported) and overall survival (p=not significant, values not reported). No analysis was undertaken for the factorial design (P-XELOX/B-XELOX combined versus P-FOLFOX-4/B-FOLFOX-4 combined).

- A pre-defined subgroup analysis on progression free survival found that the statistical superiority of bevacizumab plus chemotherapy was evident in the XELOX subgroups (B-XELOX versus P-XELOX; hazard ratio, 0.80; 97.5% CI: 0.66 to 0.96; p=not reported) but did not reach the significance level in the FOLFOX-4 subgroups (B-FOLFOX-4 versus P-FOLFOX-4; hazard ratio, 0.89; 97.5% CI: 0.74 to 1.06; p=not reported). Additional post hoc exploratory analyses (following the results from the Adjuvant Colon Cancer End Points [ACCENT] study, which found that there was a significant and direct correlation between time to recurrence after surgery and survival after recurrence in patients whose disease recurred after surgery and adjuvant treatment) showed that removing the subgroup of patients that may have slower tumour progression after adjuvant treatment (an imbalance between treatment groups with regard to an important prognostic factor which was not recognised at the start of the NO16966 trial), significantly improved (i.e. lowered) the hazard ratios for adding bevacizumab to chemotherapy compared with chemotherapy alone for both overall survival and progression free survival. Depending on the analyses conducted (e.g. exclusion of patients with prior adjuvant chemotherapy from all four treatment arms of the factorial study, or from FOLFOX groups only or from P-FOLFOX group only) the hazard ratios for overall survival ranged from 0.83 to 0.85 ($p<0.03$) and the hazard ratios for progression free survival ranged from 0.74 to 0.77 ($p<0.0001$). Although this may be plausible, the ERG note that caution should be exercised as this is a post hoc exploratory analysis.

- The majority of adverse events were generally associated with cytotoxic chemotherapy. FOLFOX-4 based regimens were generally associated with increased neutropenia/ granulocytopenia and XELOX based regimens were generally associated with increased diarrhoea and hand and foot syndrome. Adverse events that could be potentially related to bevacizumab included increased frequencies of high blood pressure, proteinuria, bleeding, gastrointestinal perforation, thromboembolic events and wound healing complications. Serious (grade 3) or life threatening (grade 4) adverse events that occurred more commonly in patients receiving bevacizumab plus chemotherapy (B-XELOX / B-FOLFOX-4 combined) than those receiving chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined) were thromboembolic events (7.8% versus 5.1%), hypertension (4.0% versus 0.8%), proteinuria (3.5% versus 0.9%) and bleeding problems (1.9% versus 1.5%), respectively. Grade 3 and 4 gastrointestinal perforations and wound healing complications were all rare (<1%). Similar results were observed when that data were restricted to the factorial analyses.

- The majority of the treatment discontinuations were attributable to chemotherapy related events rather than related to bevacizumab. Adverse events that could be potentially related to bevacizumab accounted for treatment discontinuation in 5.2% of patients in the bevacizumab plus chemotherapy group (B-XELOX / B-FOLFOX-4 combined) compared with 2.4% in the chemotherapy only (P-XELOX/ P-FOLFOX-4 combined) group (no data reported for the comparison against P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined). The statistical analysis comparing the rates of discontinuation between treatment groups were not reported in the MS or in the manufacturer's supplementary evidence.

1.3 Summary of submitted cost effectiveness evidence

The submitted cost effectiveness evidence reports on QALYs using the data from the N016966 trial. The ERG requested several changes to the modelling (including additional analyses) and a summary of the resulting incremental cost effectiveness ratios (ICER) is presented below.

Scenario	**ICERs (£ per QALY saved)**	
	B-XELOX vs. XELOX	B-FOLFOX6 vs FOLFOX6
MS original analysis		
Without APAS : Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled	£ 82,098	£ 94,989
With APAS : Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled	£ 34,170	£ 41,388
MS supplementary data (all with APAS)		
Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled	£ 35,912	£ 36,569
Analysis using the 2x2 part of N016966, XELOX and FOLFOX arms pooled	£ 48,111	£ 39,771
Analysis using 2x2 part of N016966, XELOX and FOLFOX arms unpooled	£ 35,662	£ 62,714
Analysis using the 2x2 part of N016966, XELOX and FOLFOX arms pooled, without prior adjuvant treatment	£ 36,006	£ 31,174

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

- The manufacturer conducted a limited but systematic search for clinical and cost-effectiveness studies of bevacizumab and its use in colorectal cancer. It appears unlikely that any additional trials would have met the inclusion criteria had the search been widened to include more free text terms or to include other databases.

- The NO16966 trial is of reasonable methodological quality (with some limitations), and measured a range of outcomes that are as appropriate and clinically relevant as possible.

1.4.2 Weaknesses

- The processes undertaken by the manufacturer for identifying and screening references for inclusion in the systematic review are inappropriate and the procedure applying quality criteria to included studies are not explicitly clear in the MS. These factors limit the robustness of the systematic review.

- Despite no evidence to suggest that the statistical validity of the factorial approach was methodologically inappropriate, the validity of simply pooling data from essentially two different study designs (i.e. a two arm design and a 2x2 factorial design) without accounting for between study variability is inappropriate. Unweighted (for uncertainty) pooling of results from different studies is not advisable as there are almost certainly differences between trials and which, if not accounted for, are likely to lead to biased estimates of effect. The appropriateness of combining data from the two parts of the study was also questioned by the European Medicines Agency in their assessment of extending the licensed indication of bevacizumab in combination with fluoropyrimidine-based chemotherapy for the treatment of patients with metastatic carcinoma of the colon or rectum. The resulting pooled data (manufacturer's primary pooled analysis of superiority and non inferiority) should therefore be treated with caution.

1.4.3 Areas of uncertainty

- Although it is probable that the addition of bevacizumab to oxaliplatin-based chemotherapy increases progression free and overall survival, the size of the actual treatment effect of bevacizumab is uncertain, given the trial design limitations (two part

study, open label design, imbalance of known prognostic factor [time between primary treatment and recurrence] and relatively short duration of chemotherapy treatment (aproximately 6 months) despite the fact that the trial protocol allowed coninuation of the study therapy until progressive disease or unacceptable toxicity) and the interpretation of the statistical analyses (pooled analysis of all patients versus analysis by factorial design).

- There is uncertainty around whether bevacizumab treatment should be continued until progression of the underlying disease.

- The main areas of uncertainty within the cost-effectiveness analysis relate to the choice of efficacy and health related quality of life (HRQoL) data, and the differences in treatment duration and continuity between the trial and clinical practice.

1.5 Key issues

A number of issues were identified that had an impact on the ICERs. These included the following:

- Avastin Patient Access Scheme (APAS): At the time of writing the decision on whether the proposed APAS scheme would be accepted was unknown. The majority of the analysis presented by the manufacturer included the APAS. Running the model without the APAS resulted in much higher ICERs.

- Efficacy data: It is unclear whether the clinical evidence from the randomised controlled trial used in the MS should be pooled (without weighting for uncertainty) according to data from the initial two arm part and the 2x2 factorial part of the NO16966 study or restricted to patients in the 2x2 factorial part, as per the original statistical trial plan of the NO16966 trial. Additionally it is unclear whether patients with prior adjuvant chemotherapy should be excluded from the analysis. The restriction to the trial data from the 2x2 part of the NO16966 study, the unpooling of the XELOX and FOLFOX arms, and the restriction to the data of patients without prior adjuvant chemotherapy, all have a large impact on the resulting ICERs. The restriction of the analyses to the 2x2 part of the N016966 trial increased the ICERs, exclusion of patients with prior adjuvant chemotherapy decreased the ICERs, and pooling the XELOX and FOLFOX arms affected the XELOX and FOLFOX ICERs in different directions.

- HRQoL data: The MS does not make use of the range of utility values identified from the literature review and do not explain why these values were not used. The sources of the

utility values used in the MS were poorly referenced resulting in the ERG being unable to verify them. There is also uncertainty around the clinical plausibility of the post treatment pre-progression utility value. The distributions used for the utility values in the probabilistic sensitivity analyses (PSA) reflect the uncertainty relating to the specific values used but underestimate the uncertainty relating to the selection of utility values. Using wider distributions for utility values would significantly increase the confidence intervals around the mean ICERs from the PSA. Reducing the utility values by 20% markedly increased the ICERs.

- Treatment duration: In clinical practice, treatment with non-oxaliplatin chemotherapy components may continue beyond oxaliplatin cessation although in the N016966 trial was rarely seen. Due to the structure of the Avastin Patient Access Scheme (APAS) (in which oxaliplatin is received free of charge) this could have a significant impact on the ICERs. The ERG ran an exploratory analysis to determine the effect on the ICER of stopping oxaliplatin only one month earlier and assuming incremental effectiveness is unchanged. This exploratory analysis markedly increased the ICERs.

- Intermittent versus continuous chemotherapy: Current care in England is often intermittent treatment with chemotherapy. The trial and the model both represent continuous treatment chemotherapy. The difference in cost and effectiveness between intermittent and continuous treatment is unclear. As an example, if intermittent treatment was cheaper than continuous treatment whilst having a similar efficacy, then the ICER for continuous treatment with bevacizumab versus intermittent treatment would be greater than the ICERs for continuous treatment with bevacizumab versus continuous treatment.

2. BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer's description of the underlying health problem is brief and fairly accurate. However, the manufacturer's discussion of context (p24-31, MS) lacks detail on the epidemiology (incidence and/or prevalence), aetiology, and prognosis (staging and overall survival rates) of (metastatic) colorectal cancer.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer's overview of current service provision is adequate although some discussion around specific points is required.

The MS (p25, 107-108) contains the results of the manufacturer's market research to determine the usage of chemotherapy by regimens and line (first, second or third) in the NHS in England and Wales. Limited data was provided however, the manufacturer failed to provide a detailed description of the methods undertaken for the market research. The ERG do not have experience of critically appraising market research data, and are not aware of any standard methodology doing this. The ERG's clinical advisors, however, indicated that the results of the market research data appear to be representative of first line treatment of metastatic colorectal cancer in England.

The MS (p30) suggest that patients currently receiving first line chemotherapy with fluoropyrimidine alone (5-fluorouracil plus folinic acid or capecitabine monotherapy) or fluoropyrimidine plus irinotecan are deemed unsuitable for more aggressive combination chemotherapy with oxaliplatin. However, the ERG's clinical advisors indicated that aggressive combination therapy with oxaliplatin may be used as a suitable treatment option at reduced doses or a 'stop and go' management approach may be taken.

3. CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

A summary of the decision problem addressed by the MS is shown in Table 1.

Table 1: Decision problem as issued by NICE and addressed by the MS

	Final scope issued by NICE	Decision problem addressed in the submission
Population	People with metastatic colorectal cancer for whom oxaliplatin-including chemotherapy regimens are suitable	The UK Marketing Authorisation permits bevacizumab use with oxaliplatin-based chemotherapy at any line of therapy. However, the manufacturer will be seeking a positive recommendation for these combinations in first-line only.
Intervention	Bevacizumab in combination with oxaliplatin and either 5-fluorouracil or capecitabine	Bevacizumab in combination with oxaliplatin and either 5-fluorouracil or capecitabine
Comparator(s)	• Oxaliplatin-including chemotherapy regimens without bevacizumab • Irinotecan-including chemotherapy regimens without bevacizumab	**Primary analysis** Oxaliplatin-including chemotherapy regimens without bevacizumab **Secondary analyses** Irinotecan-based regimens are considered of limited clinical relevance. However, for completeness an economic comparison has been performed versus irinotecan-based therapy, given there may be a small number of patients for whom this comparison is relevant
Outcomes	The outcome measures to be considered include: • Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life	The outcome measures considered included: • Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life
Economic Analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services

	Final scope issued by NICE	Decision problem addressed in the submission
	perspective.	perspective.
Subgroups to be considered	Guidance will only be issued in accordance with the marketing authorisation. If evidence allows the appraisal should consider the use of continuation rules based on tumour response	Consideration will be given to the activity of bevacizumab in patients with isolated liver metastases because the recent cetuximab guidance from NICE has defined this as a group where different approaches to drug therapy may be required.

3.1 Population

The manufacturer's statement of the decision problem appropriately defines the population as 'people with metastatic colorectal cancer for whom oxaliplatin-including chemotherapy regimens are suitable.' Although the term 'suitable' was not defined in the NICE scope, the MS (p26) suggest that this patient population would include people that are not resistance to adjuvant oxaliplatin (i.e. having progressed during or soon after stopping oxaliplatin-based adjuvant therapy or in those patients for whom oxaliplatin is contraindicated e.g. pre-existing neuropathy). In addition, the MS does not include any details on the mean age at diagnosis in the UK against which to compare the characteristics of patients in the clinical trial.

3.2 Intervention

Bevacizumab is a recombinant humanised monoclonal antibody that inhibits the action of vascular endothelial growth factor (VEGF), by binding to receptors on endothelial cells and thereby neutralising the physiological activity of VEGF. This reduces development of blood vessels within tumours and inhibits tumour growth.

Bevacizumab is currently licensed in the EU (including the UK) in combination with fluoropyrimidine-based chemotherapy for the treatment of patients with metastatic carcinoma of the colon or rectum. The licensed dose, administered by intravenous infusion, is either 5 mg/kg or 10 mg/kg of body weight given once every two weeks or 7.5 mg/kg or 15 mg/kg of body weight given once every three weeks. The ERG note that the clinical efficacy of the higher licensed dose has not been demonstrated in a randomised clinical trial of patients with metastatic colorectal cancer.[1]

Additional licensed indications (not the subject of this appraisal) to the products market authorisation include the following:

- bevacizumab in combination with paclitaxel or docetaxel is indicated for first-line treatment of patients with metastatic breast cancer;
- bevacizumab in addition to platinum-based chemotherapy, is indicated for first-line treatment of patients with unresectable advanced, metastatic or recurrent non-small cell lung cancer other than predominantly squamous cell histology and
- bevacizumab in combination with interferon alfa-2a is indicated for first-line treatment of patients with advanced and/or metastatic renal cell cancer.

3.3 Comparators

The decision problem addressed in the MS states that the standard comparator considered was oxaliplatin-including chemotherapy regimens without bevacizumab (primary analysis). However, the final scope issued by NICE states that comparisons should be made with (1) oxaliplatin-including chemotherapy regimens without bevacizumab (2) irinotecan-including chemotherapy regimens without bevacizumab

The manufacturer's decision to include oxaliplatin-including chemotherapy regimens without bevacizumab as the main comparator was based on evidence from market research data undertaken in December 2008, which suggests (p25, 108, MS) that oxaliplatin-based therapies are the most commonly used chemotherapy regimens for the first-line treatment of metastatic colorectal cancer in England (28% FOLFOX [intravenous 5-fluorouracil plus folinic acid plus oxaliplatin] and 24% XELOX [oral capecitabine plus intravenous oxaliplatin]). The MS (p13-16, 25, 108) states that a comparison with irinotecan-based chemotherapy regimens was considered of limited clinical and economic relevance because it is not commonly used within the UK (12% FOLFIRI [intravenous 5-fluorouracil plus folinic acid plus irinotecan] and 4% XELIRI [oral capecitabine plus irinotecan]) and is largely restricted to patients where oxaliplatin is contraindicated. In addition, approximately 25% of patients receive fluoropyrimidine monotherapy (capecitabine, 21%; intravenous 5-florouracil plus folinic acid, 4%) first-line for metastatic disease. These patients are those where the clinician and/or patient take the view that the additional toxicity conferred by oxaliplatin or irinotecan is unacceptable. The manufacturer

considered this patient population to be outside the scope of this appraisal, which is concerned with patients for whom oxaliplatin-based chemotherapy regimens would be suitable.

Although the ERG acknowledges that oxaliplatin-based chemotherapy regimens without bevacizumab are the most potentially relevant comparators for all patients with metastatic colorectal cancer, it also considers irinotecan-based chemotherapy regimens as potentially relevant comparators. The use of this treatment is also advocated by current NICE guidance which recommends FOLFIRI as a first-line treatment option for metastatic colorectal cancer.[2] While the systematic review undertaken in the MS did not consider irinotecan based chemotherapy regimens as a relevant comparator, the manufacturer undertook an economic evaluation with irinotecan-based chemotherapy (given the small number of patients for whom this comparison is relevant) for completeness.

3.4 Outcomes

The NICE scope outlines five clinical outcome measures and one measure of cost-effectiveness. All of these are stated to have been addressed in the MS (p13-15). Clinical outcome measures included overall survival, progression-free survival, response rate, adverse effects of treatment and health-related quality of life (HRQoL). These are all appropriate and clinically meaningful outcomes, and there are no other valid outcomes which the ERG would have expected to be included. Incremental cost per quality adjusted life years (QALYs) gained was used as a measure of cost-effectiveness, which is in accordance with the NICE reference case.[3]

3.5 Time frame

The manufacturer's time horizon in the health economic model was eight years. The MS (p109) states that this is equivalent to a life time horizon in the population of interest. The ERG acknowledges that the time horizon is appropriate with less than 0.1% of the population being alive at the end of eight years.

3.6 Other relevant factors

While the UK marketing authorisation does not specify a line of treatment for the use of bevacizumab in combination with fluoropyrimidine-based chemotherapy for the treatment of

patients with metastatic colorectal cancer, the manufacturer is seeking approval for first-line use only (p13, MS). Although randomised clinical trial evidence is available to support the use of bevacizumab in combination with an oxaliplatin containing regimen as a second-line therapy (relapsed disease) for metastatic colorectal cancer, the MS states (p13, 27, 106) that it is not proposing second-line use because a preliminary analysis of cost-effectiveness indicated that second-line use of bevacizumab, at the dose tested in the E3200 trial[4] (twice that proposed for first-line use) would not meet NICE's cost-effectiveness thresholds. Therefore the manufacturer would not be able to demonstrate a case for bevacizumab in this setting. The ERG acknowledges that the manufacturer has provided the wider evidence base for bevacizumab from the E3200 trial in the MS for completeness but is not the focus of the submission.

The final scope issued by NICE states that if evidence allows the appraisal should consider the use of continuation rules based on tumour response (the ERG notes that this is absent in the manufacturers statement of the decision problem). The MS (p105) states that the summary of product characteristics for bevacizumab recommends treatment to be continued until progression of the underlying disease. However, the economic analysis is based on the observed treatment duration in the NO16966 study,[5] where the average treatment duration was less than the time to progression (i.e. treatment with bevacizumab was often stopped at the same time-point as the base chemotherapy was stopped). The economic model is an accurate representation of treatment duration as it occurred in the trial but there is reason to believe that a longer duration of treatment with bevacizumab and 5-fluorouracil may be seen in clinical practice. A longer duration of treatment would significantly increase incremental costs and may also increase survival times. The ERG is unclear of the treatment effect and cost effectiveness if bevacizumab is provided as per the summary of product characteristics recommendations.

No other relevant subgroup analyses are explicitly stated in the final scope issued by NICE. However, in the manufacturers definition of the decision problem (p14, MS), it is stated that consideration will be given to the activity of bevacizumab in patients with isolated liver metastases because the NICE guidance on cetuximab defined this as a group where a different approach to drug therapy may be required.[6] Although this post-hoc subgroup analysis has been considered in the MS (p71), the ERG notes that the coverage is inadequate, as the results (including reference sources) have been poorly reported. An evaluation of the cost effectiveness of adding bevacizumab to oxaliplatin-based chemotherapy in patients with liver metastases was not undertaken by the manufacturer, despite an ERG request.

Bevacizumab is a fixed cost per cycle through the APAS. This is £800 and £1200 per 2 weekly and 3 weekly cycles respectively. The cost of bevacizumab is free after 1 year and oxaliplatin is provided free for all patient registered with the APAS receiving bevacizumab.

4. CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

4.1.1 Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

The search strategies supplied reflect a reasonable attempt to identify the literature relating to bevacizumab and its use in colorectal cancer. All the required databases have been searched (although the strategy for the Cochrane Controlled Trials Register [CCTR] is missing, and therefore was not re-run for this report). The search strategies supplied for BIOSIS, MEDLINE and EMBASE were from the Dialog system. Due to access restraints, these were re-run by the ERG in the OVID and ISI Web of Science system. Although there were some significant differences between the number of search results found by the ERG compared with those reported in the MS, in most cases this is likely due to differences in searching between the two systems and the time lag between the searches being conducted and tested. Also the ERG was unable to re-run the search on the Health Economics Evaluation Database (HEED) as the ERG do not have access to this database.

However, the search strategies seem to show some inconsistencies and are occasionally meandering, with terms appearing in the search history which are never combined or incorporated into the main search strategy and with no evidence that they have even been tested with other terms to see how their inclusion might have affected the overall result. For example in the MEDLINE Economics search strategy the terms *cost.mp, economic.mp, (health adj technology adj appraisal).mp and (colorectal adj cancer).mp* are all listed but never incorporated into the rest of the search strategy. The drug itself is not fully explored, with the alternative term *Avastin* notable by its absence. The overall appearance of the searches is one of specificity not sensitivity – for example in the MEDLINE clinical effectiveness strategy the term *bevacizumab* is combined using the boolean operator AND with the MESH term **antibodies, monoclonal/* - presumably to limit the number down to the most relevant papers but given the very small number of papers published this seems somewhat counter-productive. The searches might have benefited from a broader approach and the application of published methodological filters to identify utility/economics data.[7]

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The MS describes an inappropriate method of identifying and screening references for inclusion in the systematic review. The MS (provided as supplementary data) states that data selection and abstraction was undertaken by one individual. To ensure reproducibility and minimize selection bias assessment of eligibility of studies, and extraction of data from study reports, should be done by at least two people, independently.[8]

Details of the inclusion and exclusion criteria, as specified in the MS (p40) and that provided as supplementary data, for the systematic review of the literature is summarised in Table 2.

Table 2: Inclusion/exclusion criteria in the MS study selection

Criteria	Clinical effectiveness
Inclusion	• Population Patients with metastatic colorectal cancer • Intervention Oxaliplatin-based chemotherapy plus bevacizumab • Comparator Oxaliplatin-based chemotherapy without bevacizumab

The specified inclusion criteria are (mostly) appropriate and generally reflect the information given in the decision problem; however, there appears to be some irregularities and ambiguities in the MS.

The MS does not explicitly report any inclusion criteria relating to the study design, outcomes of interest or publication type. The ERG assumes that the review of clinical effectiveness was limited to phase III randomised controlled trials only and excluded non-English language papers and non-human studies. Although the MS included non-randomised studies (which were identified via the original searches or known to the manufacturer's information expert), it is unclear if any additional literature searches were undertaken to identify non-randomised studies. In addition, the MS did not state whether published systematic reviews and meta-analysis of primary studies would be considered in the review. The identification and assessment of such

studies would have been useful to identify any additional studies not identified by the literature searches.

The decision problem addressed in the MS states that the UK marketing authorisation permits the use of bevacizumab in combination with oxaliplatin-based chemotherapy for the treatment of patients with metastatic colorectal cancer at any line of therapy. Whilst this appears to be the approach for the systematic review in the MS (p33-73), the manufacturer is seeking approval of bevacizumab for first-line use only (Title page, p13, MS). The ERG believes the systematic review of clinical effectiveness should be a clearly defined focused review on first-line use only, with additional supportive evidence presented for other lines of therapy, for completeness.

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded

The MS identified two, head-to-head, phase III, randomised, active-controlled trials that investigated the addition of bevacizumab to an oxaliplatin-containing regimen, one as first-line therapy (patients not previously treated for their metastatic disease, NO16966 trial)[5] and one in second-line therapy (for previously treated patients with metastatic disease, E3200 trial).[4] Details of the study design and patient characteristics are summarised in Table 3.

The MS (p41-42, 84-96) also identified one non-randomised study (Three Regimens of Eloxatin Evaluation [TREE] study)[9] and two phase IV observational studies (The Bevacizumab Expanded Access Trial, BEAT[10] and the Bevacizumab Regimens: Investigation of Treatment Effects and Safety, BRiTE)[11] that provided additional data on the efficacy and safety of bevacizumab. The manufacturer (p42, MS) states that they are not aware of any relevant ongoing studies of bevacizumab combined with oxaliplatin-based chemotherapy.

Although no evidence synthesis in the form of a meta-analysis or multiple treatment comparison analyses was undertaken by the manufacturer (further discussion is provided in section 4.1.7 and 4.2.2), additional evidence from a meta-analysis of randomised controlled trials comparing the efficacy and safety of chemotherapy plus bevacizumab with chemotherapy alone in metastatic colorectal cancer[12] and a mixed treatment comparison analyses of survival and disease progression

benefits with different treatment regimens (including bevacizumab plus oxaliplatin or irinotecan for first-line and /or non-first-line treatment) for advanced colorectal cancer[13] was also identified (p73-77, MS).

The manufacturer's QUORUM diagram (provided as supplementary evidence) relating to the literature searches conforms to the QUORUM statement flow diagram (www.consort-statement.org); however, the MS does not provide a full and explicit breakdown of the reasons why all citations were rejected, especially after full text papers were retrieved for detailed evaluation (reasons for exclusion for some full text papers were reported).

Table 3: Characteristics of included studies

Study	Design	Participants (inclusion criteria)	Interventions [a,b]	Outcomes	Follow up
NO16966 trial[5]	Phase III, multicentre, multinational (32 countries including the UK), randomised, active controlled trial (Two part study - Part 1: initial two-arm open label study and Part 2 [after protocol amendment]: 2x2 factorial, double blind [for bevacizumab], active-placebo controlled trial)	Adults (male and female ≥18 years of age with Eastern Cooperative Oncology Performance Status of 0 or 1) with histologically confirmed, metastatic adenocarcinoma of the colon or rectum, not previously treated (no prior systemic therapy for advanced colorectal cancer - first-line treatment)	Part 1: Initial two arm T1': XELOX (n=317) T2': FOLFOX-4 (n=317) Part 2: 2x2 factorial, four arm T1: B-XELOX (n=350) T2: B-FOLFOX-4 (n=350) T3: P-XELOX (n=350) T4: P-FOLFOX-4 (n=351)	Co-Primary study endpoints The co-primary study endpoints after protocol modification were: • Superiority of progression free survival in patients receiving chemotherapy (B-XELOX/B-FOLFOX-4) is superior to chemotherapy alone (P-XELOX/P-FOLFOX-4) • Non-inferiority of progression free survival in patients receiving XELOX with or without bevacizumab is equivalent to FOLFOX-4 with or without bevacizumab Secondary endpoints • Progression free survival for superiority of XELOX over FOLFOX • Overall Survival • Overall Rate of Best Response • Time to Response • Duration of Response • Duration of Complete Response • Time to Treatment Failure • Safety	Median 28 months
E3200 trial[4]	Phase III, multicentre (220 sites in the USA), randomised, open-label controlled trial	Adults (male and female ≥18 years of age with Eastern Cooperative Oncology Performance Status of 0 to 2) with histologically confirmed, advanced or metastatic adenocarcinoma of the colon or rectum previously treated (second-line treatment) with fluoropyrimidine and irinotecan based regimens (either	T1: B-FOLFOX (n=293) T2: FOLFOX-4 (n=292) T3: B alone (n=244) (recruitment to T3 terminated at interim efficacy analysis after survival determined to be inferior)	Primary study endpoint • Overall survival Secondary endpoints • Response Rate • Progression free survival) • Duration of response. • Safety	Median 28 months

Study	Design	Participants (inclusion criteria)	Interventions [a,b]	Outcomes	Follow up
		separately or in combination for advanced disease)			

[a,] All treatment regimens (scheduled to receive at least 48 weeks of treatment in the NO16966 trial or) were continued until disease progression or unacceptable toxicity

[b] Individual regimens are as follows

XELOX regimen: XELOX consisted of a 2 hour intravenous infusion of oxaliplatin 130mg/m^2 on day 1 followed by oral capecitabine 1000mg/m^2 twice daily on days 1 through 14 (28 doses) of a 21 day cycle.

FOLFOX-4 regimen: FOLFOX-4 consisted of folinic acid given at a dose of 200 mg/m^2/day followed by bolus 5- fluorouracil 400 mg/m^2 and a 22 hour infusion of 5-fluorouracil 600 mg/m^2 for two consecutive days. Oxaliplatin was administered on day 1 at the dose of 85 mg/m^2 as a 2 hour infusion, concurrently with folinic acid. The treatment was repeated every 2 weeks (14 day cycle).

P-XELOX regimen: Placebo was administered as a 30 to 90 minute intravenous infusion before oxaliplatin at a dose of 7.5mg/kg on day 1 of a 21 day cycle when given with XELOX

B-XELOX regimen: Bevacizumab was administered as a 30 to 90 minute intravenous infusion before oxaliplatin at a dose of 7.5mg/kg on day 1 of a 21 day cycle when given with XELOX

P-FOLFOX-4 regimen: Placebo was administered as a 30 to 90 minute intravenous infusion before oxaliplatin at a dose of 5mg/kg on day 1 of a 14 day cycle when given with FOLFOX-4

B-FOLFOX-4 regimen: Bevacizumab was administered as a 30 to 90 minute intravenous infusion before oxaliplatin at a dose of 5mg/kg (in the NO16966 trial) or 10mg/kg (in the E3200 trial) on day 1 of a 14 day cycle when given with FOLFOX-4

B- alone: Bevacizumab was administered as a 30 to 90 minute intravenous infusion at a dose of 10mg/kg on day 1 of a 14 day cycle

4.1.4 Details of any relevant studies that were not included in the submission?

Although there were some significant differences between the repeat searches using the manufacturer's search terms compared to those reported in the MS, it is very likely (in most cases) due to differences in searching between the different systems (manufacturer used the Dialog system whereas the ERG used the OVID and ISI Web of Science systems) and the time lag between the searches being conducted and tested. The ERG is confident that all relevant studies were included in the MS and details of ongoing trials that are likely to be reporting additional evidence within 12 months were reported.

4.1.5 Description and critique of manufacturers approach to validity assessment

The validity assessment tool used in the MS is generally reflective of the quality assessment criteria developed by NICE.[14] However, it is not clear whether this was done by a single reviewer or consensus of multiple reviewers. The completed validity assessment tool for the two trials, as reported in the MS, is reproduced (with minor changes) in Table 4. The ERG acknowledges that the validity assessment tool used in the MS was appropriate.

The majority of the data for the validity assessment appears to be derived from the trial protocol (which was not requested by the ERG) and is not published in the peer reviewed articles. As a result, it was not possible for the ERG to check the validity of the manufacturer's quality assessment; however some further discussion around specific points is required.

Table 4. Validity assessment of completed trials included by the manufacturer

Validity assessment	Trials	
	Primary study NO16966	**Supportive study ECOG E3200**
How was allocation concealed?	In the assessment of bevacizumab efficacy a matched placebo was used to which patients and investigators were blind. For the comparison of oral capecitabine and intravenous 5-fluorouracil, placebo control was impractical and unethical (widespread use of intravenous placebo). Therefore, patients and clinicians were unblinded to treatment allocation. However, primary end-point was objective (tumour shrinkage on a scan) and the investigator assessment of response was checked using radiologists blind to treatment allocation	This was an open label study. However, the primary study end-point of overall survival is not liable to investigator bias
What randomisation technique was used?	Acceptable. Centralised, using interactive voice recognition system (adaptive randomisation)	Acceptable, based on limited information Centralised by the Eastern Cooperative Oncology Group Co-ordinating Centre
Was follow-up adequate?	Yes. Analyses for primary end-point (progression free survival) and overall survival was event-driven as specified in the statistical plan.	Yes. Study was stopped at a protocol specified interim analysis which demonstrated that (as specified in the trial statistical analysis plan) the O'Brien-Fleming boundary for the primary end-point had been crossed with alpha controlled at 0.00167. A final analysis for survival was subsequently conducted when 91% of FOLFOX and 89% of B-FOLFOX patients had died with a median follow-up of 25.0 and 28 months, respectively.
Were the individuals undertaking the outcomes assessment aware of allocation?	The primary analysis was based on investigator assessment of progression free survival. Investigators were blinded to treatment allocation of bevacizumab or placebo, but not to the allocation of XELOX versus FOLFOX. A supportive analysis conducted by independent reviewers blind to all treatment allocation was conducted.	No, but this was irrelevant to the primary end-point in this study (overall survival)
Was the RCT conducted in the UK (or were one or more centres of the multinational RCT located in the UK)? If not, where was the RCT conducted, and is clinical practice likely to differ from	This was a multinational study conducted by 216 investigators from 32 countries including the UK. The principal investigator on the study was Prof James Cassidy from the Beatson Oncology Centre in Glasgow. Clearly Prof Cassidy felt that the	No. This study was conducted in the USA. The main difference between the study population in the USA and the UK is probably in the first-line treatment that they received. In the UK, the predominant first-line chemotherapy is oxaliplatin plus a

Validity assessment	Trials	
	Primary study NO16966	**Supportive study ECOG E3200**
UK practice?	protocol was relevant and appropriate for UK clinicians and patients.	fluoropyrimidine. In the USA at the time of the study it was irinotecan plus 5-fluorouracil plus folinic acid making it logical to examine the role of FOLFOX+/-B in the second-line setting. Thus data from this second-line study are being used to support UK use in the first-line setting. This is acceptable given the general view, supported by NICE guidance, that the sequence of oxaliplatin and irinotecan-based chemotherapies for the first two lines of treatment for metastatic colorectal cancer is unimportant.
How do those included in the RCT participants compare with patients who are likely to receive the intervention in the UK? Consider factors known to affect outcomes in the main indication, such as demographics, epidemiology, disease severity, setting	The patient population represents one that is relevant from a UK perspective – patients with metastatic colorectal cancer, with a slight excess of males over females receiving their first treatment for metastatic colorectal cancer with a combination of a fluoropyrimidine and oxaliplatin (the predominant first-line treatment in the UK). The obvious difference from the general population of UK patients diagnosed with metastatic colorectal cancer is that they are slightly younger – 83% of patients diagnosed with colorectal cancer in the UK are over 60 years of age, whereas the mean age of recruits to the present study is 59.7 years. However, this is probably more typical of the fitter, younger patients who would receive combination chemotherapy with oxaliplatin-based combinations in this country.	
For pharmaceuticals, what dosage regimens were used in the RCT? Are they within those detailed in the Summary of Product Characteristics?	Yes. The dose of 7.5 mg/kg bevacizumab every three weeks accords with the summary of product characteristics dose range. Doses of cytotoxic agents used accord with the relevant summary of product characteristics and UK clinical practice.	Yes. The dose of 10 mg/kg bevacizumab every two weeks accords with the summary of product characteristics dose range. Doses of cytotoxic agents used accord with the relevant summary of product characteristics and UK clinical practice.
Were the study groups comparable?	Yes.	Yes
Were the statistical analyses used appropriate?	Yes. The studies were analysed in accordance with the predetermined statistical plan prepared by statisticians.	Yes. The studies were analysed in accordance with the predetermined statistical plan prepared by statisticians.

Validity assessment	Trials	
	Primary study NO16966	**Supportive study ECOG E3200**
Was an intention-to-treat analysis undertaken?	Yes – primary superiority analyses (bevacizumab versus placebo) were done on an intention to treat basis with primary non-inferiority analyses (XELOX versus FOLFOX) done on a per protocol population basis as is appropriate	Yes – primary superiority analyses (bevacizumab versus placebo) were done on an intention to treat basis
Were there any confounding factors that may attenuate the interpretation of the results of the RCT(s)?	None known	None known

RCT, randomised controlled trial; XELOX, oral capecitabine plus intravenous oxaliplatin; FOLFOX, intravenous 5-fluorouracil plus folinic acid plus oxaliplatin; B-FOLFOX, FOLFOX plus bevacizumab; B-XELOX, XELOX plus bevacizumab

The MS states (p45, 48-49) that randomisation was performed centrally with stratification (the ERG assumes this was performed by a computer) and treatment assignments were concealed (central randomisation system). However, the manufacturer's validity assessment (p63, MS) suggests that randomisation assignment was not concealed. The ERG notes that there has been some confusion in the MS between the terms allocation concealment and open label design. The ERG acknowledges that adequate methods of randomisation and allocation concealment were used in the two trials; however, patients and investigators were all unblinded (open label design) to the assigned treatment in the N16966 trial (except for bevacizumab therapy which was double blinded) and the E3200 trial. Double blinding protects against performance bias and measurement bias and it absence from randomised controlled trial tends to result in larger treatment effects. Double blinding safeguards against performance bias and measurement bias[15] and its deficiency in randomised controlled trials tends to result in larger treatment effects.[16] With many cytotoxic cancer drugs, the nature of the interventions prohibits blinding (i.e. drug toxicities or manner of administration) for the practical and ethical reason that informed dose monitoring and adjustment is required. Although it is almost universally absent from oncology trials, blinded outcome assessment can amplify bias reduction.[17]

The MS states (p64) that the demographic characteristics of the patients in the NO16966 trial were generally representative of the characteristics expected of this population in the UK, albeit younger and fitter. This observation is not unusual since there is a tendency for randomised controlled trial participants to be younger and fitter people than those who might be treated in routine clinical practice. The performance status of participants in randomised controlled trials and real-life populations is also favourable (Eastern Cooperative Oncology Group performance status ≤ 1 in the NO16966 trial and Eastern Cooperative Oncology Group performance status PS ≤ 2 in the E3200 trial).

The MS states (p64) that the study groups were well balanced at baseline in the NO16966 trial and the E3200 study; however, it is unclear whether or not the manufacturer formally looked for statistically significant differences between treatment groups (p53-54, MS).

Although the demographic and prognostic data were well balanced between the initial two-arm part and the 2x2 factorial part of the NO16966 trial, the ERG notes the following exceptions. The number of Caucasian patients that enrolled in the 2x2 factorial part of the study were approximately 10% greater than the initial two-arm part of the study. The percentages of patients with an ECOG performance status of 0 in the 2x2 factorial part of the study were approximately 10% higher compared with patients in the initial two-arm part of the study. Although previous adjuvant treatment, an important prognostic characteristic, was well balanced between the two parts of the study, the Adjuvant Colon Cancer End Points (ACCENT) study (a collection of individual patient data from 18 trials testing fluorouracil-based adjuvant therapy for patients with stage II and III colon cancer [n=17381], but data-analysis restricted to a subset of patients [n=5722] who experienced tumour recurrence after initial therapy) showed that there was a significant ($p<0.0001$) and direct correlation between time to recurrence after surgery and survival after recurrence in patients whose disease recurred after surgery and adjuvant treatment. These data suggested that tumours that do not recur for long periods of time following initial treatment of the primary disease tend to behave in a more indolent fashion following tumour recurrence (i.e. patients with slower tumour growth leading to later recurrence after surgery have an improved survival due to slow growth of the recurrent tumour).[18] In the NO16966 trial, the median time from end of adjuvant treatment to randomisation in the initial two-arm part of the study (FOLFOX, 517 days; XELOX, 511 days) was shorter than the 2x2 factorial part of the trial (B-FOLFOX, 623 days; B-XELOX, 597 days, P-FOLFOX, 769 days; P-XELOX, 660 days). More notably, the P-FOLFOX group had the lowest proportion of patients with a time from start of adjuvant treatment to randomisation of less than one year and the highest proportion of patients with a time to recurrence≥4 years than any of the other treatment groups . The ERG acknowledges that this imbalance could not have been accounted for in the NO16966 trial, which was initiated prior to the findings from the ACCENT data set. Overall, it is not clear how these disparities may have influenced or biased the results.

In the NO16966 trial and the E3200 trial, less than 20% of participants in each treatment group were reported to have been loss to follow-up. In general, the greater the number of subjects who are lost,

the more the trial may be subject to bias because patients who are lost often have different prognoses from those who are retained. Patients may discontinue their participation in studies because they are not prepared to accept the treatment, they recover and move address or because they have died.[19] In both trials, all patients were accounted for and efficacy analysis was conducted using the intention-to-treat approach (p56-57, MS). Overall, attrition bias should be low in these studies.

4.1.6 Description and critique of manufacturers outcome selection

As discussed in Section 3.4, the ERG considers the manufacturer's outcome selection to be relevant and appropriate. The outcome measures described in the decision problem generally reflect those in the NO16966 trial and E3200 study and include overall survival, progression-free survival, response rate, health-related quality of life (not assessed in either trial) and adverse effects of treatment.

4.1.7 Describe and critique the statistical approach used

The manufacturer did not undertake a meta-analysis (p73-74, MS). The ERG notes that while the systematic review in the MS is based on the inclusion of two large pivotal studies, one in first-line treatment (NO16966 study) and the other in second-line treatment (E3200 trial), the manufacturer is seeking approval for first-line use only (and is the basis of the MS). In addition, due to the differences in the study populations (demographics, baseline disease characteristics and bevacizumab dosages) a meta-analysis of the NO16966 trial and E3200 study would not be appropriate.

The statistical analysis of the NO16966 trial was adequately reported by the manufacturer; however, the ERG believes that the validity of simply pooling data from the initial two-arm part and the 2x2 factorial part of the study (which was planned a priori and allowed for in the sample size power calculations) without accounting for between study variability is inappropriate. The NO16966 trial was originally designed to demonstrate that XELOX was non superior (equivalence trial) to FOLFOX-4. After publication of studies demonstrating the benefit of adding bevacizumab to irinotecan plus 5-fluorouracil plus folinic acid,[20,21] the protocol of the original NO16966 study design was amended to a 2x2 factorial randomised phase II trial to address an additional primary objective of superiority of progression free survival of bevacizumab in combination of chemotherapy (B-XELOX or B-FOLFOX) versus chemotherapy alone (P-XELOX or P-FOLFOX). As these are essentially two different study designs (a two arm design and a 2x2 factorial design) it is unclear to the ERG how between study variability in the estimate for the baseline treatment mean for patients receiving

chemotherapy alone (XELOX or FOLFOX) with those receiving XELOX or FOLFOX plus placebo is accounted for and how randomisation has been preserved in the two study designs (two arm design and 2x2 factorial design) when estimating population treatment effects. Unweighted (for uncertainty) pooling of results from different studies is not advisable as there are almost certainly differences between trials (as noted in section 4.1.5) and which, if not accounted for, are likely to lead to biased estimates of effect.[22,23,24]

Moreover, the supplementary data provided by the manufacturer (p7) states that an exploratory analysis based on all patients from the initial two-arm part and the 2x2 factorial part of the NO16966 trial was included in the analysis plan (for regulatory approval to the European Medicines Agency [EMEA] in Europe), to be used in case of borderline results for progression free survival in the primary analysis of superiority of bevacizumab in combination of chemotherapy (B-XELOX or B-FOLFOX) versus chemotherapy alone (P-XELOX or P-FOLFOX). As the results were not borderline for the superiority analysis, the EMEA assessment report for bevacizumab[25] also questioned the appropriateness of combining the two parts of the study.

4.1.8 Summary statement

Although the majority of the MS reflects the UK marketing authorisation, it does not reflect the broader population outlined in the licensed indication and final scope issued by NICE. The licensed indication permits the use of bevacizumab in combination with fluoropyrimidine-based chemotherapy for the treatment of patients with metastatic colorectal cancer but does not specify a line of treatment (first-line or second-line). Whilst the NICE scope broadly reflects the licensed indication; the manufacturer is seeking approval for first-line use only.

The manufacturers search strategy was adequately reported and the submission appears to contain all relevant head-to-head randomised controlled trials. Although oxaliplatin-including chemotherapy regimens without bevacizumab is the most potentially relevant comparator for all patients with metastatic colorectal cancer for whom oxaliplatin-including chemotherapy regimens are suitable, no considerations or comparisons with irinotecan-including chemotherapy regimens were undertaken for the systematic review (even though an economic comparison was performed). The validity assessment tool used to assess the included studies was satisfactory, although details of process, in terms of whether it was performed by two independent reviewers, are missing. The outcomes selected were

relevant and the statistical methods well described. The submitted evidence adequately reflects the decision problem (for which approval is being sought), although there appears to be small inconsistencies which probably relate more to reporting than actual differences.

4.2 Summary of submitted evidence

Two randomised controlled trials were identified in the manufacturer's systematic review. The NO16966 trial (which forms the main pivotal evidence for the MS) was a phase III, multicentre, multinational, two-arm, randomised, open label study with the primary objective of confirming the non-inferiority of XELOX compared with FOLFOX-4 in adult patients with histologically confirmed metastatic colorectal cancer not previously treated (first-line therapy). Following randomisation of 634 patients, the open label study was amended to include a 2x2 factorial randomised (partially blinded for bevacizumab) phase III trial (n=1401) with the co-primary objective of demonstrating superiority of bevacizumab in combination with chemotherapy (B-XELOX or B-FOLFOX-4) compared with placebo (P-XELOX or P-FOLFOX-4). The dose of bevacizumab was 5 mg/kg every two weeks (B-FOLFOX-4) or 7.5 mg/kg every three weeks (B-XELOX).[5] The E3200 study (which provides supportive evidence for the efficacy and safety of bevacizumab) was a phase III, multicentre, three-arm, randomised, open label study which compared the safety and efficacy of B-FOLFOX-4 versus FOLFOX-4 versus bevacizumab alone in 829 adult patients with advanced or metastatic colorectal cancer previously treated with a fluoropyrimidine based and irinotecan based chemotherapy regimen (second-line therapy). The dose of bevacizumab was 10 mg/kg every two weeks.[4]

Further supportive evidence was provided from one non-randomised study (the TREE study)[9] and two large observational registry studies in first-line metastatic colorectal cancer (the First BEAT study[10] and the BRiTE study).[11] The non-randomised TREE study evaluated the safety, tolerability, and efficacy of three oxaliplatin regimens (bolus and infusion fluorouracil and folinic acid with oxaliplatin, bolus fluorouracil and low-dose folinic acid with oxaliplatin, or capecitabine with oxaliplatin) without (TREE-1; n=147) or with (TREE-2; n=213) bevacizumab (5 mg/kg every two weeks or 7.5 mg/kg every three weeks) as first-line treatment of patients with histologically documented metastatic or recurrent colorectal cancer and no prior treatment for advanced disease.[9] The BRITE study, conducted in the USA, treated elderly patients with any fluoropyrimidine-based chemotherapy (at the investigators' discretion) plus bevacizumab (n=1953],[11] and the First BEAT study, conducted in Europe and Canada, treated elderly patients with fluoropyrimidine-based chemotherapy (single-agent fluoropyrimidine or fluoropyrimidine plus oxaliplatin or irinotecan) at the clinician's discretion plus

bevacizumab (5 mg/kg every two weeks for 5-fluoruracil based regimens or 7.5 mg/kg every three weeks for capecitabine-based regimens); n=1914]).[10]

4.2.1 Summary of results

This section presents the main clinical evidence from the NO16966 trial (the manufacturer is seeking approval for first-line use only and this trial is therefore the basis of the MS), as reported by the manufacturer and constructed (data re-tabulated in a consistent and more transparent format) by the ERG. A tabulated summary of such data is presented in Table 5 and Table 6 (post-hoc treatment subgroup comparisons are provided in Appendix 1). Supplementary data from the E3200 trial, which provides evidence for second-line use, are also presented. Additional information (e.g. additional results and exploratory analyses), not reported in the MS, was provided by the manufacturer in the clarifications of questions raised by the ERG.

Overall survival

The manufacturers' primary pooled analysis of superiority in the NO16966 trial, which combined all patients in the trial (patients in the initial two-arm part plus patients in the 2x2 factorial part of the study) showed that the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly improved overall survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined) in patients not previously treated for their metastatic disease. For the intention to treat population, the hazard ratio for death was 0.83 (97.5% CI: 0.74 to 0.93; p=0.0019) at a median follow-up of 28 months corresponding to a 17% relative reduction in overall mortality and an increase in median overall survival from 18.9 months in the chemotherapy group to 21.2 months in the bevacizumab plus chemotherapy group (absolute difference 2.3 months).

A secondary pooled analysis of superiority, restricted to patients in the second 2x2 part of the NO16966 (as per the original statistical trial plan), showed that the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) improved overall survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4 combined), although this was not statistically significant. For the intention to treat population, the hazard ratio for death was 0.89 (97.5% CI: 0.76 to 1.03; p=0.0769) at a median follow-up of 28 months corresponding to a 11% relative reduction in overall mortality and an increase in median overall survival from 19.9 months in the chemotherapy group to 21.3 months in the bevacizumab plus chemotherapy group (absolute difference 1.4 months).

The manufacturers' pooled analysis of non-inferiority (using the eligible patient population [i.e. the intention to treat population minus patients who did not receive at least one dose of study drug, and those patients who violated major inclusion/exclusion criteria] and the intention to treat population), which pooled all patients in the NO16966 trial (patients in the initial two arm part plus patients in the 2x2 factorial part of the study) showed that the XELOX (XELOX/P-XELOX/B-XELOX) and FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4) based regimens were equivalent for overall survival (Table 5).

Table 5. Summary of overall survival from the NO16966 trial

Interventions (Regimens)[a]	Median follow up (months)	Numbers followed in each group (n)	Number of patients with event (n)	Median progression free survival, (months)	Hazard Ratio (97.5% CI; p-value)
Initial 2 arm design					
XELOX	-	317	250 (78.9%)	18.8	Not applicable
FOLFOX-4 (control)	-	317	262 (82.6%)	17.7	Not applicable
2x2 factorial design					
B-XELOX	-	350	211 (60.3%)	21.4	Not applicable
P-XELOX	-	350	231 (66.0%)	19.2	Not applicable
B-FOLFOX-4	-	349	209 (59.9%)	21.2	Not applicable
P-FOLFOX-4 (control)	-	351	224 (63.8%)	20.4	Not applicable
Manufacturers primary analysis (pooled results from both parts of study- all six groups)					
Superiority - Intention to treat analysis					
B-XELOX / B-FOLFOX-4 combined vs.	28	699 vs.	420 (60.1%) vs.	21.2 vs.	0.83 (0.74, 0.93; p=0.0019)
P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined		1335	967 (72.4%)	18.9	
Non inferiority[b] - Eligible patient population analysis					
XELOX/P-XELOX/B-XELOX combined vs.	-	NR	NR	19.7 vs.	1.00 (0.88, 1.13; p=NR)
FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4 combined				19.5	
Non inferiority[b] - Intention to treat analysis	-				
XELOX/P-XELOX/B-XELOX combined vs.		NR	NR	19.8 vs.	0.99 (0.88, 1.12; p=NR)
FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4 combined				19.6	
Manufacturers secondary analysis (analysis restricted to the 2 by 2 factorial design) [c]					
Superiority - Intention to treat analysis					
B-XELOX / B-FOLFOX-4 combined vs.	28	699 vs.	420 (60.1%) vs.	21.3 vs.	0.89 (0.76, 1.03; p=0.0769)
P-XELOX/ P-FOLFOX combined		701	455 (64.9%)	19.9	

NR, not reported; CI, confidence intervals

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4, B- alone, Bevacizumab only

[b] Non-inferiority was concluded if the upper limit of the 97.5% confidence interval of the hazard ratio was ≤ 1.23

[c] Test of the hypotheses of no interaction for overall survival between the different treatment components (FOLFOX-4, XELOX, bevacizumab, non bevacizumab) was 0.9380, which does not meet the conventional level of significance (not reported but assumed by ERG) of less than 0.05. Therefore, marginal analysis appropriate.

Table 6. Summary of progression free survival from the NO16966 trial

Interventions (Regimens)[a]	Median follow up (months)	Numbers followed in each group (n)	Number of patients with event (n)	Median progression free survival, (months)	Hazard Ratio (97.5% CI; p-value)
Initial 2 arm design					
XELOX	-	317	290 (91.5%)	7.1	Not applicable
FOLFOX-4 (control)	-	317	299 (94.3%)	7.7	Not applicable
2x2 factorial design					
B-XELOX	-	350	295 (84.3%)	9.3	Not applicable
P-XELOX	-	350	301 (86.0%)	7.4	Not applicable
B-FOLFOX-4	-	349	299 (85.7%)	9.4	Not applicable
P-FOLFOX-4 (control)	-	351	321 (91.5%)	8.6	Not applicable
Manufacturers primary analysis (pooled results from both parts of study – all six groups)					
Superiority - Intention to treat analysis					
B-XELOX / B-FOLFOX-4 combined vs. P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined	28	699 vs. 1335	594 (85.0%) vs. 1211 (90.7%)	9.4 vs. 7.7	0.79 (0.72, 0.87; p=0.0001)
Non inferiority[b] - Eligible patient population analysis					
XELOX/P-XELOX/B-XELOX combined vs. FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4 combined		NR	NR	8.0 vs. 8.5	1.02 (0.92, 1.14; p= NR)[d]
Non inferiority[b] - Intention to treat analysis					
XELOX/P-XELOX/B-XELOX combined vs. FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4 combined		NR	NR	8.0 vs. 8.5	1.01 (0.91,1.12; p=NR)[e]
Manufacturers secondary analysis (analysis restricted to the 2 by 2 factorial design)[c]					
Superiority - Intention to treat analysis B-XELOX / B-FOLFOX-4 combined vs. P-XELOX/ P-FOLFOX combined		699 vs. 701	513 vs. 547	9.4 vs. 8.0	0.83 (0.72, 0.95; p=0.0023)

NR, not reported; CI, confidence intervals

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4, B- alone, Bevacizumab only

[b] Non-inferiority was concluded if the upper limit of the 97.5% confidence interval of the hazard ratio was ≤ 1.23

[c] Test of the hypotheses of no interaction for progression free survival the different treatment components (FOLFOX-4, XELOX, bevacizumab, non bevacizumab) was 0.7025, which does not meet the conventional level of significance (not reported but assumed by ERG) of less than 0.05 . Therefore, marginal analysis appropriate.

[d] values are different to that report in the original published paper – Hazard ratio, 1.05 (0.94,1.18)[26]

[e] values are different to that report in the original published paper – Hazard ratio, 1.04 (0.93,1.16)[26]

Supplementary data from the E3200 trial (which only compared FOLFOX-4 based regimens with/without bevacizumab) showed that addition of bevacizumab to FOLFOX-4 chemotherapy significantly improved overall survival compared with FOLFOX-4 alone for previously treated patients with metastatic disease. For the intention to treat population, the hazard ratio for death was 0.751 (95% CI: 0.332 to 0.893; p=0.0012) at a median follow-up of 28 months corresponding to a 24.9% relative reduction in overall mortality and an increase in median overall survival from 10.8 months in the FOLFOX-4 group to 13.0 months (value reported as 12.9 months in original published paper)[4] in the bevacizumab plus FOLFOX-4 group (absolute difference 2.2 months).

Progression-free survival

The manufacturers' primary pooled analysis of superiority in the NO16966 trial, which combined all patients in the trial (patients in the initial two-arm part plus patients in the 2x2 factorial part of the study) showed that the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly enhanced progression free survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined). For the intention to treat population, the hazard ratio for remaining free of disease progression was 0.79 (97.5% CI: 0.72 to 0.87; p=0.0001) at a median follow-up of 28 months corresponding to a 21% relative reduction in disease progression or death and an increase in median progression free survival from 7.7 months in the chemotherapy group to 9.4 months in the bevacizumab plus chemotherapy group (absolute difference 1.7 months).

A secondary pooled analysis of superiority, restricted to patients in the second 2x2 part of the NO16966 (as per the original statistical trial plan) showed that the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly improved progression free survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4 combined). For the intention to treat population, the hazard ratio for remaining free of disease progression was 0.83 (97.5% CI: 0.72 to 0.95; p=0.0023) at a median follow-up of 28 months corresponding to a 17% relative reduction in disease progression or death and an increase in median progression free survival from 8.0 months in the chemotherapy group to 9.4 months in the bevacizumab plus chemotherapy group (absolute difference 1.4 months)

The manufacturers' pooled analysis of non-inferiority (using the eligible patient population and the intention to treat population) showed that the XELOX (XELOX/P-XELOX/B-XELOX) and FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4) based regimens were equivalent for progression free survival (Table 6).

Supplementary data from the E3200 trial showed that addition of bevacizumab to FOLFOX-4 chemotherapy significantly improved progression free survival compared with FOLFOX-4 alone for previously treated patients with metastatic disease. For the intention to treat population, the hazard ratio for remaining free of disease progression was 0.518 (97.5% CI: 0.416 to 0.646; $p<0.0001$) at a median follow-up of 28 months (value reported as 0.61 [CI not reported] in the original published paper)[4] corresponding to a 48.2% relative reduction in disease progression or death and an increase in median progression free survival from 4.5 months in the chemotherapy group to 7.5 months in the bevacizumab plus chemotherapy group (values reported as 4.7 and 7.3 months respectively, in the original published paper)[4] with an absolute difference of 3 months.

Response rates

The data for the response rates were not reported in the original MS for the NO16966 trial; the manufacturer's supplementary evidence, which included data on response rates, was poorly reported and incomplete. For example, the response rates of only three study arms (B-XELOX, B-FOLFOX and P-FOLFOX) were reported instead of six (as requested by the ERG). The ERG notes that the original published paper by Saltz et al.,[5] which reported the results from the 2x2 factorial part of the NO16966, found that the response rates for the intention to treat population were similar between the bevacizumab plus chemotherapy group (B-XELOX / B-FOLFOX-4 combined, 47%) and chemotherapy alone (P-XELOX/ P-FOLFOX-4 combined; 49%) group (odds ratio, 0.90; 97.5% CI: 0.71 to 1.14; p=0.31). Similarly, the original published paper by Cassidy et al.,[26] which combined all patients in the trial (patients in the initial two-arm part plus patients in the 2x2 factorial part of the study) showed that the overall response rates for all patients treated with XELOX (XELOX/P-XELOX/B-XELOX combined, 47%) compared with FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/B-FOLFOX-4 combined, 48%) were similar (odds ratio, 0.94; 97.5% CI: 0.77 to 1.15; p=not reported).

Subgroup analysis

A post hoc subgroup analysis on the impact of adding bevacizumab to oxaliplatin-based chemotherapy in patients with liver metastases in the NO16966 trial could not be tabulated due to the poor and restrictive reporting of the limited results in the MS (p71) (additional information requested by the ERG was not provided), thus a narrative summary, as reported in the MS, is provided. However, these data should be interpreted with caution.

The NO16966 study showed that the addition of bevacizumab to chemotherapy (it was not clear if the analysis included all patients from the NO16966 trial or restricted to the 2x2 factorial part of the study) appears to improve both R0 resection rates and outcomes after resection. The R0 (i.e. removal of metastasis/ses with a margin of healthy tissue) resection rate was 6.3% for patients receiving bevacizumab plus chemotherapy compared with 4.9% for patients receiving chemotherapy alone. Although the resection rate difference did not reach statistical significance, the 2 year survival (it was not clear if this is overall survival or progression free survival) increased from 82.3% (95% CI: 69.4 to 95.1) in the chemotherapy group to 90.9% (95% CI: 82.4 to 99.4) in the bevacizumab plus chemotherapy group (p-value not reported).

While the overall data from the NO16966 trial showed a progression free survival benefit for those patients treated with bevacizumab plus oxaliplatin-based chemotherapy, a pre-defined subgroup analysis based on the type of chemotherapy received in the 2x2 part of the study showed that a benefit was only statistically significant for the XELOX groups (hazard ratio, 0.80; 97.5% CI: 0.66 to 0.96; p=0.0059; however, in the original published paper the hazard ratio is reported as 0.77; 97.5% CI: 0.63 to 0.94; p=0.0026) but not the FOLFOX group (hazard ratio, 0.89; 97.5% CI: 0.74 to 1.06; p=0.1312; however, in original published paper the hazard ratio is reported as 0.89; 97.5% CI: 0.73 to 1.08; p=0.1871). An exploratory analyses undertaken by the manufacturer (provided as supplementary data) found that patients in the FOLFOX groups who had received prior adjuvant therapy did not derive benefit from bevacizumab (hazard ratio, 1.75; 97.5% CI: 1.15 to 2.65; p=not reported), whereas patients who did not have adjuvant therapy did derive a benefit (hazard ratio, 0.72; 97.5% CI: 0.58 to 0.90; p=not reported). One potential explanation for this discrepancy, as noted in section 4.1.5 and section 5.3, is that the P-FOLFOX-4 group had a longer time interval between the end of adjuvant chemotherapy to relapse compared with the other groups (FOLFOX, 517 days; XELOX, 511 days [initial two-arm study] ; B-FOLFOX, 623 days; B-XELOX, 597 days, P-FOLFOX, 769 days; P-XELOX, 660 days [2x2

factorial study]) despite stratification for key prognostic factors suggesting there was an imbalance due to the better prognosis in some patients in this group. This led to an unusually high progression free survival in the P-FOLFOX subgroup of patients (8.6 months) compared with the original FOLFOX-4 group (7.7 months). This imbalance was not present in the P-XELOX group (7.4 months compared with 7.1 months in the original XELOX group) or in the other study groups. Additional exploratory analyses (provided as supplementary evidence by the manufacturer) showed that removing the subgroup of patients that may have slower tumour progression, significantly improved the hazard ratios for adding bevacizumab to chemotherapy compared with chemotherapy alone for both overall and progression free survival. Depending on the analyses conducted (e.g. exclusion of patients with prior adjuvant chemotherapy from all four treatment arms of the factorial study, or from FOLFOX groups only or from P-FOLFOX group only; Appendix 2) the hazard ratios for overall survival ranged from 0.83 to 0.85 ($p<0.03$) and the hazard ratios for progression free survival ranged from 0.74 to 0.77 ($p<0.0001$). Although this may be plausible, the ERG note that caution should be exercised as this is a post hoc exploratory analysis.

Critique of efficacy data reported

There are a number of issues that may limit the robustness of the efficacy data reported in the MS. The main issue relates to the methodological limitations of the pooled statistical analysis, undertaken by the manufacturer, which has been described and critiqued in section 4.1.7.

Although the FOLFOX-4 regimen was used in the in the NO16966 trial, the MS (p21, 102-103 and supplementary evidence) states the vast majority of patients in England and Wales receive the FOLFOX-6 regimen. In addition, whilst there have been no direct comparisons between FOLFOX-4 and FOLFOX-6, these regimens are most widely used in clinical trials and in clinical practice. FOLFOX-6 is considered to offer equivalent efficacy to FOLFOX-4 whilst being less resource intensive to deliver (patients are only required to attend hospital for infusions once per 2 weekly cycle as opposed to twice with the FOLFOX-4 regimen). The ERG and their clinical advisors have no reasons to dispute the statements made by the manufacturer. The ERGs clinical advisors also indicated that it is generally accepted that the FOLFOX-4 regimen offers equivalent clinical outcomes to the XELOX regimen.

The manufacturer's statistical analyses of the NO16966 trial based their interpretation on the upper limit of a 1-sided 97.5% confidence interval, which is the same as the upper limit of a 2-sided 95% confidence interval. Although both 1-sided and 2-sided confidence intervals allow for inferences about non-inferiority, the guidelines for reporting of non-inferiority and equivalence randomised trials recommend 2-sided confidence interval reporting as more appropriate in non-inferiority trials.[27]

Although the addition of bevacizumab to oxaliplatin-based chemotherapy significantly improved overall survival compared with chemotherapy alone in the manufacturers primary pooled analysis (which combined all patients in the trial), no significant improvement was observed when the analysis was restricted to patients in the second 2x2 part of the study (as per the original statistical trial plan) only. Despite this difference (which may be partly explained by the imbalance of patients that may have slower tumour progression as noted earlier), and the lack of statistical power to assess this endpoint, a gain in overall survival is not commonly observed in an era where three subsequent treatment lines have demonstrated benefits in overall survival as compared with best supportive care. Wagner et al.[28] suggest that when considering the relatively short treatment duration of bevacizumab in first-line therapy (median approximately 6 months in the NO16966 trial), as well as the number of subsequent therapy lines with additional impact on survival, a benefit in overall survival is neither a sensitive, nor realistic endpoint for a first-line therapy study.

Although a statistically significant treatment action was ruled out in the 2x2 factorial part of the NO16966 study (p=0.7025), a high p-value could reflect low power and so cannot be taken as evidence for no interaction.[29] However, based on the clarifications received from the manufacturer, the ERG is confident that the NO16966 trial was adequately powered for the pooled non-inferiority and superiority comparisons, which were event driven.

Data checking the MS highlighted numerous errors and inconsistencies between the MS and peer reviewed published papers of the N016966 trial and the E3200 trial (some of which have already been highlighted). For example, the MS reports the median duration of follow up in the NO16966 study (initial two-arm design and the 2x2 factorial design) as 28 months; whereas the original peer reviewed paper suggest 29.7 months.[26] Similarly, the MS reports the median

duration of follow up for the 2x2 factorial part of the NO16966 trial as 28 months; whereas the original peer reviewed paper suggest 27.6 months.[5]

Safety and tolerability

This section presents the main safety evidence from the NO16966 trial,[5] as reported by the manufacturer and constructed (data re-tabulated in a consistent and more transparent format) by the ERG. Supplementary information, not reported in the MS, was provided by the manufacturer in the clarifications of questions raised by the ERG. Although very limited, additional safety data were reported from one non-randomised study[9] and two phase IV observational studies.[10,11] No additional safety data were reported from the E3200 trial.[4] The MS (p78) states that the safety data from the E3200 trial (which compared B-FOLFOX-4 versus FOLFOX-4 versus bevacizumab alone) is of limited value as it provides no information on the B-XELOX regimen and the study employed a dose of bevacizumab twice that proposed for use in clinical practice. The ERG notes that the provision of this information would have been useful.

The analysis of adverse events in the MS is based primarily on the comparison of the pooled bevacizumab containing groups (B-XELOX / B-FOLFOX-4 combined) with no bevacizumab containing groups (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined). An additional analysis (requested by the ERG), restricted to patients in the second 2x2 factorial part (B-XELOX / B-FOLFOX-4 combined versus P-XELOX/ P-FOLFOX-4 combined) of the NO16966 trial (as per the original statistical trial plan and reported in the primary published peer reviewed clinical paper) was also presented.

A summary of the rates of discontinuation, including reasons for premature termination are presented in Table 7 and 8 (data for each treatment group are provided in Appendix 3 and 4). It is noteworthy, that the manufacturer failed to provide full details of adverse events leading to treatment discontinuation in the original MS. Those that were provided subsequently, at the request of the ERG, were incomplete. Although the rates of discontinuation in the NO16966 trial were higher in the bevacizumab containing groups (B-XELOX/ B-FOLFOX-4 combined, 30.8%) than in the no bevacizumab containing groups (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined, 25.3%), the statistical analysis comparing the rates of discontinuation between the treatment groups were not reported in the MS or in the requested supplementary data.

Corresponding data, restricted to the 2x2 factorial analyses, yielded similar results (B-XELOX/ B-FOLFOX-4 combined, 30.8% versus P-XELOX/ P-FOLFOX-4 combined, 20.8%, respectively). The manufacturer's supplementary evidence states that the majority of the treatment discontinuations were attributable to chemotherapy related events rather than related to bevacizumab (i.e. neurotoxicity, gastrointestinal events, and hematologic events). Events that could be potentially related to bevacizumab accounted for treatment discontinuation in 5.2% and 2.4% of patients in the bevacizumab (B-XELOX / B-FOLFOX-4 combined) and chemotherapy only groups (P-XELOX/ P-FOLFOX-4 combined), respectively (no data reported for the comparison against P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined). Analysis of treatment withdrawals in the 2x2 factorial part of the study, showed that despite protocol allowance of treatment continuation until disease progression, only 29% (203/699) and 47% (329/701) of bevacizumab and chemotherapy alone recipients, respectively, were treated until progression indicating that a large proportion of patients stopped treatment earlier than allowed by the study protocol. In general, the rates of discontinuation were similar between the FOLFOX-4 containing regimens and XELOX containing regimens

In the NO16966 trial, nearly all patients in each treatment group (99.0 to 99.7%) experienced at least one adverse event (all grades). The overall incidence of the most commonly occurring adverse events was similar between the bevacizumab plus chemotherapy group (B-XELOX / B-FOLFOX-4 combined, 99.4%) compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined, 99.2%). Increases in the incidence (absolute 5% difference) of common adverse events with the addition of bevacizumab to chemotherapy compared with chemotherapy alone included the following: stomatitis (35.1% versus 29.3%), hand-foot syndrome (27.1% versus 20.8%), bleeding problems (30.5% versus 23.6%) and hypertension (19.0% versus 4.4%), respectively. Similar results were observed when that data were restricted to the 2x2 factorial analyses.

Table 7. Reasons for stopping treatment (by type of pooled analysis) during the primary treatment phase of the NO16966 trial (data derived from manufacturer's supplementary evidence)

Reasons for stopping treatment	Manufacturers primary pooled analysis of all six groups combined (n)				Manufacturers secondary 2x2 factorial analysis (n)			
	Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4		Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4	
	B-XELOX / B-FOLFOX-4 combined	XELOX / FOLFOX-4 / P-XELOX / P-FOLFOX-4 combined	XELOX / P-XELOX / B-XELOX combined	FOLFOX-4 / P-FOLFOX-4 / B-FOLFOX-4 combined	B-XELOX / B-FOLFOX-4 combined	P-XELOX / P-FOLFOX-4 combined	P-XELOX / B-XELOX combined	P-FOLFOX-4 / B-FOLFOX-4 combined
Subjects randomised (intention to treat population)	699	1335	1017	1017	699	701	700	700
Safety	**226 (32.3%)**	**363 (27.2%)**	304 (29.9%)	285 (28.0%)	**226 (32.3%)**	**151 (21.5%)** [f]	191 (27.3%)	186 (26.6%)
Abnormality of lab test	0	0	0	0	0	0	0	0
Adverse event [b]	**210 (30.0%)**	**334 (25.0%)**	280 (27.5%)	264 (26.0%)	**210 (30.0%)**	**144 (20.5%)** [g]	181 (25.9%)	173 (24.7%)
Death	16 (2.3%)	29 (2.2%)	24 (2.4%)	21 (2.1%)	16 (2.3%)	7 (1.0%)	10 (1.4%)	13 (1.9%)
Non-safety	**367 (52.5%)** [d]	**857 (64.2%)**	596 (58.6%)	628 (61.8%)	**367 (52.5%)**	**472 (67.3%)**	414 (59.1%)	425 (60.7%)
Insufficient therapeutic response	**203 (29.0%)**	**587 (44.0%)**	406 (39.9%)	384 (37.8%)	**203 (29.0%)**	**329 (46.9%)**	275 (39.3%)	257 (36.7%)
Early improvement	0	0	0	0	0	0	0	0
Violation of selection criteria at entry	6 (0.9%)	9 (0.7%)	5 (0.5%)	10 (1.0%)	6 (0.9%)	7 (1.0%)	5 (0.7%)	8 (1.1%)
Other protocol violation	1 (0.1%)	3 (0.2%)	1 (0.1%)	3 (0.3%)	1 (0.1%)	2 (0.3%)	1 (0.1%)	2 (0.3%)
Refused treatment [c]	65 (9.3%) [e]	105 (7.9%)	71 (7.0%)	99 (9.7%)	65 (9.3%)	58 (8.3%)	54 (7.7%)	69 (9.9%)
Failure to return	2 (0.3%)	10 (0.7%)	5 (0.5%)	7 (0.7%)	2 (0.3%)	0	0	2 (0.3%)
Other	90 (12.9%)	143 (10.7%)	108 (10.6%)	125 (12.3%)	90 (12.9%)	76 (10.8%)	79 (11.3%)	87 (12.4%)
Total	**593 (84.8%)**	**1220 (91.4%)**	900 (88.5%)	913 (89.8%)	593 (84.8%)	623 (88.9%)	605 (86.4%)	611 (87.3%)

Note: Frequencies in **BOLD** indicate an absolute 5% difference (higher or lower) between bevacizumab with/without chemotherapy or XELOX and FOLFOX-4

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4

[b] Includes intercurrent illness

[c] Including "did not co-operate", "withdrew consent"

[d] value reported as 366 (52.4%) in original peer reviewed published paper[5]

[e] value reported as 64 (9.2%) in original peer reviewed published paper[5]

[f] value reported as 150 (21.4%) in original peer reviewed published paper[5]

[g] value reported as 143 (20.4%) in original peer reviewed published paper[5]

Table 8. Number of patients discontinuing treatment (by type of pooled analysis) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

	Manufacturers primary pooled analysis of all six groups combined (n)				Manufacturers secondary 2x2 factorial analysis (n)			
	Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4		Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4	
	B-XELOX / B-FOLFOX-4 combined	XELOX / FOLFOX-4 / P-XELOX / P-FOLFOX-4 combined	XELOX / P-XELOX / B-XELOX combined	FOLFOX-4 / P-FOLFOX-4 / B-FOLFOX-4 combined	B-XELOX / B-FOLFOX-4 combined	P-XELOX / P-FOLFOX-4 combined	P-XELOX / B-XELOX combined	P-FOLFOX-4 / B-FOLFOX-4 combined
Number of patients (safety population)	**695**	**1303**	**1008**	**990**	**695**	**674**	**692**	**677**
Discontinued treatment due to adverse event								
All grade	**214 (30.8%)**	**330 (25.3%)**	**280 (27.8%)**	**264 (26.7%)**	**214 (30.8%) [b]**	**140 (20.8%) [b]**	**181 (26.2%)**	**173 (25.6%)**
Grade 1 and 2 only	**NR**	**NR**	**NR**	**NR**	**62 (8.9%) [c]**	**40 (5.9%) [c]**	**NR**	**NR**
Grade 3 and 4 only	**NR**	**NR**	**NR**	**NR**	**145 (20.9%) [c]**	**101 (15.0%) [c]**	**NR**	**NR**
Bevacizumab targeted	**NR**	**NR**	**NR**	**NR**	**36 (5.2%)**	**16 (2.4%)**	**NR**	**NR**
Treatment-related deaths	**15 (2.2%)**	**30 (2.3%)**	**24 (2.4%)**	**21 (2.1%)**	**15 (2.2%)**	**13 (1.9%)**	**1.4 (2.0%)**	**1.4 (2.1)**

Note: Frequencies in BOLD indicate an absolute 5% difference (higher or lower) between bevacizumab with/without chemotherapy or XELOX and FOLFOX-4
NR, not reported
[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4
[b] values from the MS (as presented in the above table) are different to that reported in the original paper[5]: B-XELOX / B-FOLFOX-4 combined, 207 (30%) versus P-XELOX/ P-FOLFOX-4 combined, 141 (21%)
[c] these numbers were not reported in the MS and have been derived from the original paper[5], therefore, these numbers do not sum to the value for 'all grade'; however, they do sum to the value reported in note [b] (see above)

A summary of the pooled serious (Grade 3) and life threatening (Grade 4) adverse events, from the NO16966 study, are summarised in Table 9 (data for each treatment group are provided in Appendix 5). In general, the overall incidence of the serious and life threatening adverse events were higher in the bevacizumab plus chemotherapy group (B-XELOX / B-FOLFOX-4 combined, 79.9%) compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined, 74.8%). Increases in the incidence of grade 3 and 4 adverse events with the addition of bevacizumab to chemotherapy compared with chemotherapy alone included the following: stomatitis, diarrhoea, nausea/vomiting, gastrointestinal perforation, bleeding problems, hand-foot syndrome, venous thromboembolic events, arterial thromboembolic events, hypertension, proteinuria and cardiac disorders**.** Similar results were observed in the 2x2 factorial analyses. The MS states (p79) that the majority of these adverse events are generally associated with cytotoxic chemotherapy and that the increased incidence is likely to be a consequence of longer chemotherapy treatment duration amongst bevacizumab recipients (treatment was continued until disease progression and the time to disease progression was increased by the inclusion of bevacizumab in the treatment regimen).

Adverse events of special interest to bevacizumab included hypertension, proteinuria, bleeding, gastrointestinal perforation, thromboembolic events and wound healing complications. The most common of these in the NO16966 trial was thromboembolic events. The occurrence of grade 3 and 4 hypertension, proteinuria and bleeding was 1.9 to 4%. Grade 3 and 4 gastrointestinal perforations and wound healing complications were all rare (<1%). Similar results were observed when that data were restricted to the 2x2 factorial analyses.

In general, the overall incidence of the serious and life threatening adverse events were higher in the FOLFOX-4 containing regimens (FOLFOX-4/ P-FOLFOX-4/ B-FOLFOX-4 combined, 80.3%) compared with XELOX containing regimens (XELOX/ P-XELOX/ B-XELOX combined, 72.8%). Grade 3 and 4 adverse events that were more common (absolute 5% difference) for XELOX based regimens than FOLFOX-4 based regimens included diarrhoea (20.8% versus 11.9%) and hand and foot syndrome (8.1% versus 1.4%), respectively. In contrast, grade 3 and 4 adverse events that were more common (absolute 5% difference) for FOLFOX-4 based regimens than XELOX based regimens included blood and lymphatic disorders (mainly neutropenia / granulocytopenia, 48.2% versus 22.7%, respectively). Similar results were observed when that data were restricted to the 2x2 factorial analyses.

Table 9. Grade 3 and 4 adverse events (by type of pooled analysis) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

	Manufacturers primary pooled analysis of all six groups combined (n)				Manufacturers secondary 2x2 factorial analysis (n)			
	Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4		Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4	
	B-XELOX / B-FOLFOX-4 combined	XELOX / FOLFOX-4 / P-XELOX / P-FOLFOX-4 combined	XELOX / P-XELOX / B-XELOX combined	FOLFOX-4 / P-FOLFOX-4 / B-FOLFOX-4 combined	B-XELOX / B-FOLFOX-4 combined	P-XELOX / P-FOLFOX-4 combined	P-XELOX / B-XELOX combined	P-FOLFOX-4/ B-FOLFOX-4 combined
Number of patients (safety population)	695	1303	1008	990	695	674	692	677
All grade 3 and 4	555 (79.9%)	974 (74.8%)	734 (72.8%)	795 (80.3%)	555 (79.9%)	503 (74.6%)	503 (72.7%)	555 (82.0%)
Grade 4 only [b]	164 (23.6%)	242 (18.6%)	141 (14.0%)	265 (26.8%)	164 (23.6%)	118 (17.5%)	96 (13.9%)	186 (27.5%)
Any related serious adverse event	182 (26.2%)	288 (22.1%)	238 (23.6%)	232 (23.4%)	182 (26.2%)	149 (22.1%)	174 (25.1%)	157 (23.2%)
Gastrointestinal disorders								
All grade 3 and 4	219 (31.5%)	383 (29.4%)	348 (34.5%)[c]	254 (25.7%)	219 (31.5%)	186 (27.6%)	246 (35.5%)	159 (23.5%)
Grade 4 only	18 (2.6%)	34 (2.6%)	35 (3.5%)	17 (1.7%)	18 (2.6%)	16 (2.4%)	23 (3.3%)	11 (1.6%)
Stomatitis	19 (2.7%)	21 (1.6%)	15 (1.5%)	25 (2.5%)	19 (2.7%)	12 (1.8%)	13 (1.9%)	18 (2.7%)
Diarrhoea	121 (17.4%)	207 (15.9%)	210 (20.8%)	118 (11.9%)	121 (17.4%)	104 (15.4%)	147 (21.2%)	78 (11.5%)
Nausea/vomiting	63 (9.1%)	99 (7.6%)	90 (8.9%)	72 (7.3%)	63 (9.1%)	42 (6.2%)	66 (9.5%)	39 (5.8%)
Blood and lymphatic disorders								
All grade 3/4	284 (40.9%)[d]	422 (32.4%)	229 (22.7%)	477 (48.2%)	284 (40.9%)[d]	214 (31.8%)	173 (25.0%)	325 (48.0%)
Grade 4 only	117 (16.8%)	138 (10.6%)	58 (5.8%)	197 (19.9%)	117 (16.8%)	74 (11.0%)	52 (7.5%)	139 (20.5%)
Neutropenia/ granulocytopenia	163 (23.5%)	328 (25.2%)	71 (7.0%)	420 (42.4%)	163 (23.5%)	178 (26.4%)	51 (7.4%)	290 (42.8%)
Febrile neutropenia	19 (2.7%)	37 (2.8%)	10 (1.0%)	46 (4.6%)	19 (2.7%)	17 (2.5%)	5 (0.7%)	31 (4.6%)
Gastrointestinal perforation	4 (0.6%)	4 (0.3%)	6 (0.6%)	2 (0.2%)	4 (0.6%)	2 (0.3%)	5 (0.7%)	1 (0.1%)
Bleeding problems	13 (1.9%)	20 (1.5%)	19 (1.9%)	14 (1.4%)	13 (1.9%)	8 (1.2%)	12 (1.7%)	9 (1.3%)

Table 9. Grade 3 and 4 adverse events (by type of pooled analysis) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

	Manufacturers primary pooled analysis of all six groups combined (n)				Manufacturers secondary 2x2 factorial analysis (n)			
	Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4		Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4	
	B-XELOX / B-FOLFOX-4 combined	XELOX / FOLFOX-4 / P-XELOX / P-FOLFOX-4 combined	XELOX / P-XELOX / B-XELOX combined	FOLFOX-4 / P-FOLFOX-4 / B-FOLFOX-4 combined	B-XELOX / B-FOLFOX-4 combined	P-XELOX / P-FOLFOX-4 combined	P-XELOX / B-XELOX combined	P-FOLFOX-4/ B-FOLFOX-4 combined
Neurological and other toxicity								
Hand/foot syndrome (grade 3)	48 (6.9%)	48 (3.7%)	82 (8.1%)	14 (1.4%)	48 (6.9%)	23 (3.4%)	61 (8.8%)	10 (1.5%)
Neurotoxicity	125 (18.0%)	221 (17.0%)	178 (17.7%)	168 (17.0%)	125 (18.0%)	130 (19.3%)	127 (18.4%)	128 (18.9%)
Venous thromboembolic events	54 (7.8%)	66 (5.1%)	47 (4.7%)	73 (7.4%)	54 (7.8%)	33 (4.9%)	31 (4.5%)	56 (8.3%)
Arterial thromboembolic events	12 (1.7%)	12 (0.9%)	11 (1.1%)	13 (1.3%)	12 (1.7%)	7 (1.0%)	10 (1.4%)	9 (1.3%)
Hypertension	28 (4.0%)	10 (0.8%)	21 (2.1%)	17 (1.7%)	28 (4.0%)	8 (1.2%)	20 (2.9%)	16 (2.4%)
Proteinuria	24 (3.5%)[e]	12 (0.9%)	33 (3.3%)[f]	3 (0.3%)	24 (3.5%)	0	21 (3.0%)	3 (0.4%)
Wound healing complications	3 (0.4%)	4 (0.3%)	3 (0.3%)	4 (0.4%)	3 (0.4%)	2 (0.3%)	3 (0.4%)	2 (0.3%)
Fistula/intrabdominal abscess	6 (0.9%)	9 (0.7%)	6 (0.6%)	9 (0.9%)	6 (0.9%)	1 (0.1%)	3 (0.4%)	4 (0.6%)
Cardiac disorders	37 (5.3%)	15 (1.2%)	20 (2.0%)	32 (3.2%)	37 (5.3%)	3 (0.4%)	16 (2.3%)	24 (3.5%)
Infections/infestations	51 (7.3%)	111 (8.5%)	66 (6.5%)	96 (9.7%)	51 (7.3%)	50 (7.4%)	40 (5.8%)	61 (9.0%)
Laborotory abnormalities								
Low neutrophils	170 (24.5%)	341 (26.2%)	83 (8.2%)	428 (43.2%)	170 (24.5%)	182 (27.0%)	53 (7.7%)	299 (44.2%)
Low haemoglobin	17 (2.4%)	35 (2.7%)	26 (2.6%)	26 (2.6%)	17 (2.4%)	13 (1.9%)	15 (2.2%)	15 (2.2%)
Low platelets	27 (3.9%)	82 (6.3%)	69 (6.8%)	40 (4.0%)	27 (3.9%)	38 (5.6%)	39 (5.6%)	26 (3.8%)

Note: Frequencies in BOLD are 5% higher or lower in absolute terms than for chemotherapy without bevacizumab

[a] **XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4**

[b] **defined as life-threatening adverse events**

Table 9. Grade 3 and 4 adverse events (by type of pooled analysis) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

	Manufacturers primary pooled analysis of all six groups combined (n)				Manufacturers secondary 2x2 factorial analysis (n)			
	Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4		Superiority analysis: Bevacizumab plus chemotherapy versus chemotherapy alone		Non – inferiority analysis: XELOX versus FOLFOX-4	
	B-XELOX / B-FOLFOX-4 combined	XELOX / FOLFOX-4 / P-XELOX / P-FOLFOX-4 combined	XELOX / P-XELOX / B-XELOX combined	FOLFOX-4 / P-FOLFOX-4 / B-FOLFOX-4 combined	B-XELOX / B-FOLFOX-4 combined	P-XELOX / P-FOLFOX-4 combined	P-XELOX / B-XELOX combined	P-FOLFOX-4/ B-FOLFOX-4 combined

[c] value reported as 342 (33.9%) in original MS
[d] value reported as 203 (29.2%) in original MS
[e] value reported as 4 (0.6%) in original MS
[f] value reported as 13 (1.3%) in original MS

Additional data (only reported as a very brief narrative summary in the MS) from one non-randomised study[9] and two phase IV observational studies[10,11] also suggest that bevacizumab is generally well tolerated.

Critique of safety data reported

The reporting and interpretation of the safety and tolerability data is generally good. Although all grade adverse events were recorded in the NO16966 trial, it is unclear in the MS which toxicity grading criteria was used. Information from the original peer reviewed published papers[26] suggest that the grading criteria was defined according to the National Cancer Institute Common Toxicity Criteria. The MS (including supplementary data) also failed to report p-values for all comparisons and the additional supportive evidence from the non-randomised and observational studies was poorly reported (i.e. all adverse events were not reported).

As noted in section 4.1.7, the ERG believes that the validity of simply pooling data from the initial two-arm part and the 2x2 factorial part of the study without accounting for between study variability is inappropriate. It is noteworthy, that for some adverse event comparisons the difference between the bevacizumab plus chemotherapy regimen and the chemotherapy alone regimen is slightly different depending on how the treatment regimens have been pooled. For example, the absolute difference between the bevacizumab plus chemotherapy regimen and the chemotherapy alone regimen in the 2x2 factorial analyses is slightly greater for certain adverse events compared with all groups combined (e.g. all grade 4 adverse events, absolute difference 6.1% versus 5.0%; diarrhoea, absolute difference 2.0% versus 1.5%; nausea/vomiting, absolute difference 2.9% versus 1.5%; treatment discontinuation due to adverse events, absolute difference 10% versus 5.5% respectively). On the other hand, it is slightly lower for other adverse events e.g. febrile neutropenia, absolute difference 0.2% versus 0.1%; neurotoxicity, absolute difference 1.3% versus 1.0% respectively.

Although the study protocol of the NO16966 trial allowed for discontinuation of just oxaliplatin in patients who experienced oxaliplatin associated toxicity whilst continuing bevacizumab, a large proportion of patients discontinued all therapy, including bevacizumab, in response to adverse events that were probably not bevacizumab related. The manufacturer's supplementary evidence suggests that as a result there may have been a greater therapeutic benefit of bevacizumab had these patients remained on treatment.

Data checking also highlighted some errors in the reporting of data between (and within) the MS, supplementary data reported by the manufacturer, and the peer reviewed published papers of the N016966 trial,[5] these have already been highlighted in Tables 7 to 9.

4.2.2 Critique of submitted evidence syntheses

No evidence synthesis in the form of a meta-analysis or mixed treatment comparison (MTC) analyses was undertaken by the manufacturer. As noted in section 4.1.7 and section 5.3.3, a meta-analysis of the NO16966 trial and E3200 study would not be appropriate.

4.2.3 Summary

The MS probably contains unbiased estimates of the treatment effect of bevacizumab (in combination with oxaliplatin-based chemotherapy) within the stated scope of the decision problem (for first-line therapy). This is based on the results of a single RCT which is of reasonable methodological quality when judged using the NICE quality assessment criteria,[14] but the reporting of the trial results is neither totally transparent nor are all results fully tabulated for each outcome. It is difficult to interpret the data with full confidence due to the trial design limitations (two part study, open label design, imbalance of known prognostic factor [time between primary treatment and recurrence] and relatively short duration of chemotherapy treatment (aproximately 6 months) despite the fact that the trial protocol allowed coninuation of the study therapy until progressive disease or unacceptable toxicity) and the complexity in interpretation of the statistical analyses. Despite no evidence to suggest that the validity of the factorial approach was methodologically inappropriate, the validity of simply pooling data from essentially two different study designs (i.e. two arm design and 2x2 factorial design) is questionable. These factors make it difficult to accurately assess the true size of the treatment effect.

The manufacturers' primary pooled analysis of superiority (using the intention to treat population) in the NO16966 trial showed that after a median follow up of 28 months, the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly improved progression free survival and overall survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined) in adult patients with histologically confirmed

metastatic colorectal cancer not previously treated (median progression free survival: 9.4 versus 7.7 months [absolute difference, 1.7 months]; hazard ratio, 0.79; p=0.0001; median overall survival: 21.2 versus 18.9 months [absolute difference, 2.3 months]; hazard ratio, 0.83; p=0.0019).

A secondary pooled analysis of superiority, restricted to patients in the second 2x2 part of the NO16966 study (as per the original statistical trial plan [B-XELOX / B-FOLFOX-4 combined versus P-XELOX/ P-FOLFOX-4 combined] and which the ERG believe to be more appropriate) also found similar results (median progression free survival: 9.4 versus 8.0 months [absolute difference, 1.4 months]; hazard ratio, 0.83; p=0.0023; median overall survival: 21.3 versus 19.9 months [absolute difference, 1.4 months]; hazard ratio, 0.89; p=0.0769).

The manufacturers' pooled analysis of non-inferiority (using the eligible patient population and the intention to treat population) showed that the XELOX (XELOX/P-XELOX/B-XELOX combined) and FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/ B-FOLFOX-4 combined) based regimens were equivalent for both progression free survival (p=not significant, values not reported) and overall survival (p=not significant, values not reported). No analyses were undertaken for the factorial design (P-XELOX/B-XELOX combined versus P-FOLFOX-4/B-FOLFOX-4 combined).

A pre-defined subgroup analysis on progression free survival found that the statistical superiority of bevacizumab plus chemotherapy was evident in the XELOX subgroups (B-XELOX versus P-XELOX; hazard ratio, 0.80; 97.5% CI: 0.66 to 0.96; p=not reported) but did not reach the significance level in the FOLFOX-4 subgroups (B-FOLFOX-4 versus P-FOLFOX-4; hazard ratio, 0.89; 97.5% CI: 0.74 to 1.06; p=not reported). Additional post hoc exploratory analyses (following the results from the Adjuvant Colon Cancer End Points [ACCENT] study, which found that there was a significant and direct correlation between time to recurrence after surgery and survival after recurrence in patients whose disease recurred after surgery and adjuvant treatment) showed that removing the subgroup of patients that may have slower tumour progression after adjuvant treatment (an imbalance between treatment groups with regard to an important prognostic factor which was not recognised at the start of the NO16966 trial), significantly improved the hazard ratios for adding bevacizumab to chemotherapy compared with chemotherapy alone for both overall survival and progression free survival. Depending on the

analyses conducted (e.g. exclusion of patients with adjuvant chemotherapy from all four treatment arms of the factorial study, or from FOLFOX groups only or from P-FOLFOX group only) the hazard ratios for overall survival ranged from 0.83 to 0.85 ($p<0.03$) and the hazard ratios for progression free survival ranged from 0.74 to 0.77 ($p<0.0001$). Although this may be plausible, the ERG note that caution should be exercised as this is a post hoc exploratory analysis

The adverse event profile of bevacizumab appears to be generally well tolerated. However, it is unclear whether bevacizumab treatment should be continued until progression of the underlying disease.

5. ECONOMIC EVALUATION

5.1 Overview of manufacturer's economic evaluation

Four versions of the model were received by the ERG:

1) Original version (with APAS)

2) Original version (without APAS)

3) Version post clarification with changes suggested by ERG (with APAS)

4) Version excluding patients with prior adjuvant therapy (with APAS)

The ERG notes that without painstaking checking the models could be slightly different in terms of data population or structure. Relevant results from each of these models will be presented.

The main changes made to the model after the ERG clarification letter (model 1 compared with model 3) were:

- Parametric curves were fitted using all PFS and OS data (rather than truncating the data at 28 months)
- PFS was modelled by fitting a Weibull curve to the survival data from 6 months onwards.
- The capability to incorporate different treatment durations for different chemotherapy components was included. Specifically oxaliplatin stopping prior to capecitabine / 5-FU was accounted for.
- Oxaliplatin wastage was included.

In addition the following options were included in the model as a result of the ERG clarification letter:

- An option to only use the data from the 2x2 part of the NO16966 trial only was included
- An option to run an analysis in which the XELOX and FOLFOX arms are not pooled was included.
- A version of the model using data from patients without prior adjuvant therapy from the 2x2 part of the NO16966 trial was included

5.2 Natural history, model structure, accuracy, discounting and timeframe

The natural history of patients receiving first-line oxaliplatin-based therapy for metastatic colorectal cancer (CRC) is modelled assuming transition between the following health states: first-line treatment, after first-line treatment but pre-progression, progressed and dead. The MS uses a cohort model with a Markov structure with a cycle length of one month and a time frame of eight years which seems appropriate. A discount rate of 3.5% per annum was used, in accordance with the NICE reference case.[3] The discount rate was applied monthly (0.29% per month) with a discount applied in month one onwards.

5.3 Treatment effectiveness

5.3.1 Results from Study NO16966

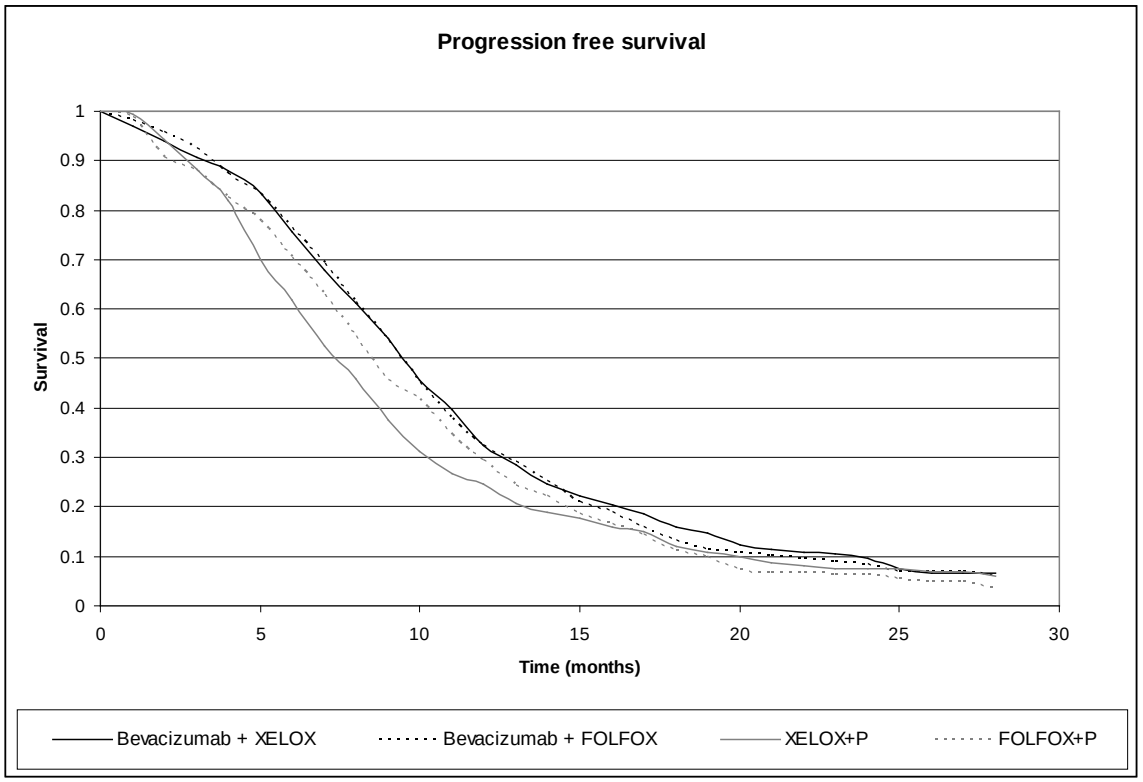
The MS base-case uses data from the intention to treat (ITT) population from the NO16966 trial pooling the initial two-arm study and the 2x2 factorial part of the trial as the base-case (thus pooling six arms in total). As discussed in Section 4.1.7 the heterogeneity between the trials was not accounted for when arms from the initial two-arm study were pooled with arms from the 2x2 factorial study.

The ERG group requested an analysis using data from the 2x2 part of the trial as presented in the original peer reviewed published paper by Saltz et al.,[5] and this was provided by the manufacturer. Using the data from only the 2x2 part of the trial resulted in longer survival times in the comparator arms, which corresponded to markedly higher ICERs for adding bevacizumab.

In the MS the XELOX and FOLFOX arms were pooled. The true relative risk of adding bevacizumab may differ when added to XELOX rather than FOLFOX, also the underlying efficacy of XELOX and FOLFOX may be different. The ERG requested an analysis in which the XELOX and FOLFOX arms were not pooled and this was provided by the manufacturer. The survival curves with XELOX and FOLFOX unpooled are presented in Figure 1 and also presented separately on p117 of the MS supplementary data. When these arms are not pooled this results in a markedly lower ICER for adding bevacizumab to XELOX and a markedly higher ICER for adding bevacizumab to FOLFOX. The reason for this can clearly be seen in Figure 1 which shows that the FOLFOX and B-FOLFOX survival curves are quite close.

The survival (PFS and OS) seen in the P-XELOX/P-FOLFOX arms of the 2x2 trial was better than that in the initial two-arm study (XELOX/FOLFOX) of the NO16966. The manufacturers suggested that this is due to an imbalance in baseline risk factors favouring the P-FOLFOX group and it is hypothesised that this is due to an imbalance in time elapsed since prior adjuvant therapy between groups; see Section 4.1.5. A correlation between time elapsed since prior adjuvant therapy and rate of tumour growth is clinically plausible. The MS presents evidence of a relationship between time since prior adjuvant therapy and survival. Figure 3 which was presented in the MS supplementary data p25 shows the distribution of patients according to time from adjuvant treatment to randomisation by treatment arm. The exclusion of patients with prior adjuvant chemotherapy from all four treatment arms reduces the hazard ratios relating to the addition of bevacizumab from 0.89 to 0.83 for OS and from 0.83 to 0.74 for PFS (reducing the number of patients from 1400 to 1060). Figure 1 shows the survival curves when the prior adjuvant chemotherapy patients are excluded and a gap between the with and without bevacizumab curves is seen. Following an ERG request the manufacturer's provided an analysis in which patients who received prior adjuvant therapy were excluded from the data set and this resulted in markedly lower ICERs.

Figure 1 Survival of each of the four arms from the 2x2 part of the N016966 trial



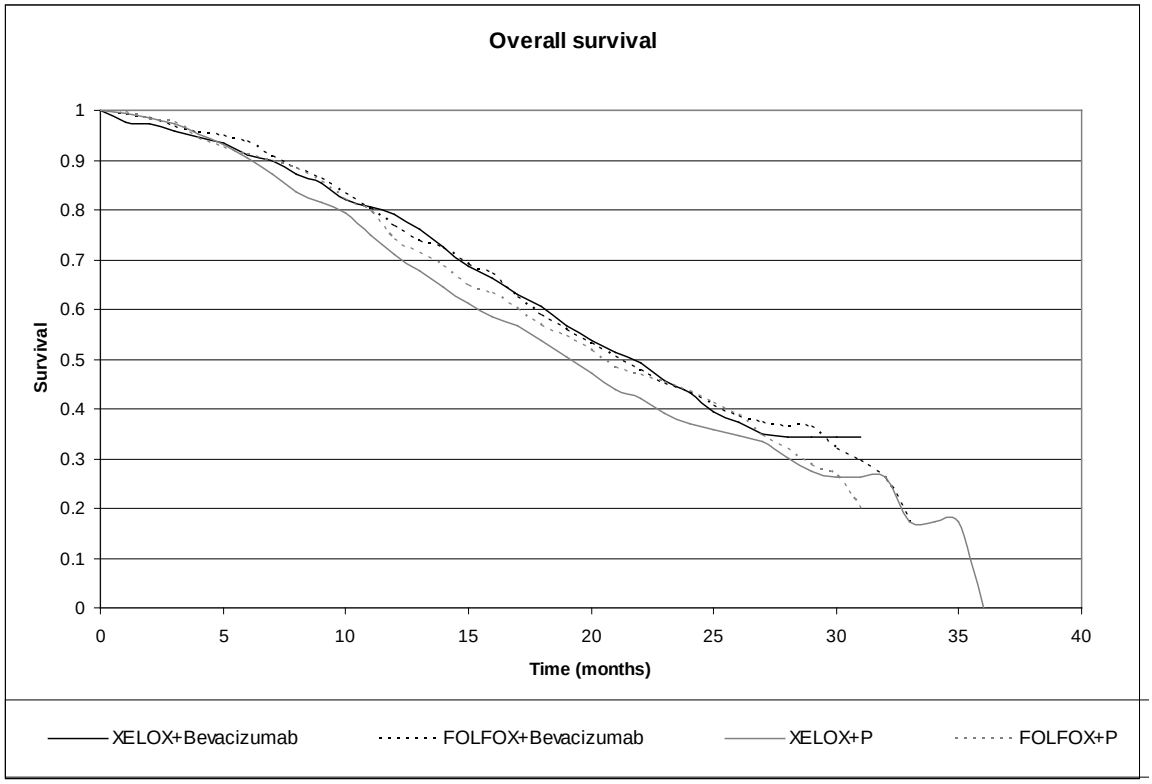
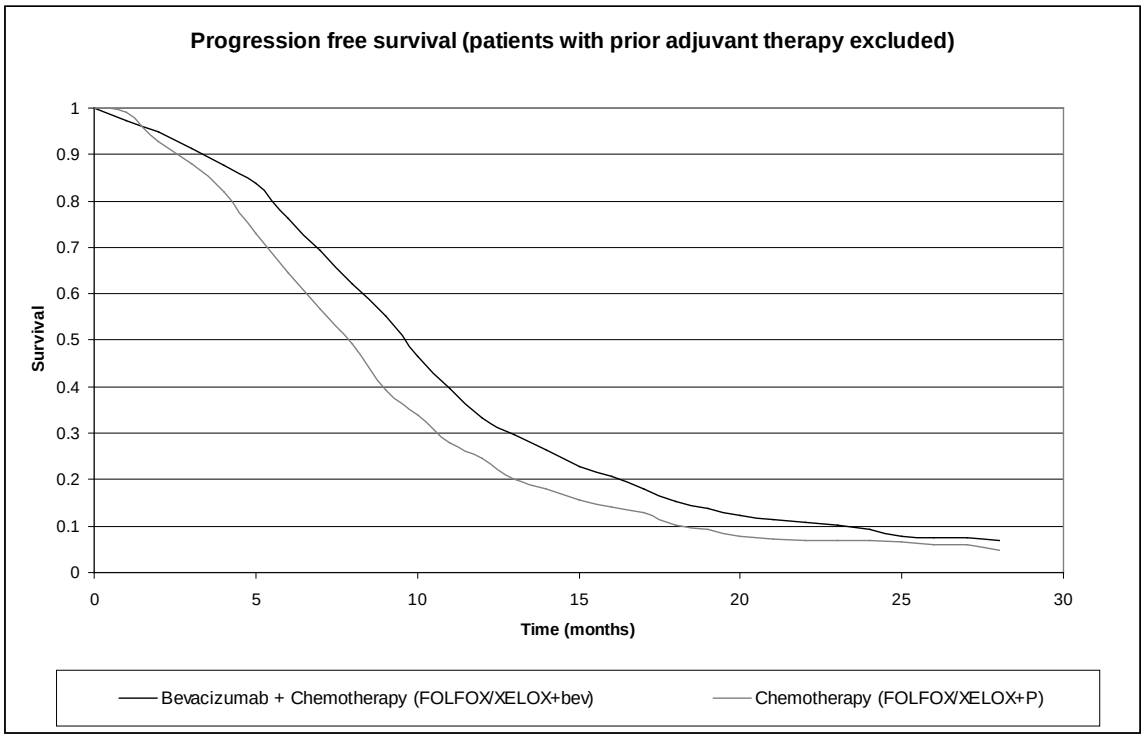


Figure 2 **Survival curves when patients with prior adjuvant therapy are excluded from the 2x2 part of the N016966 trial**



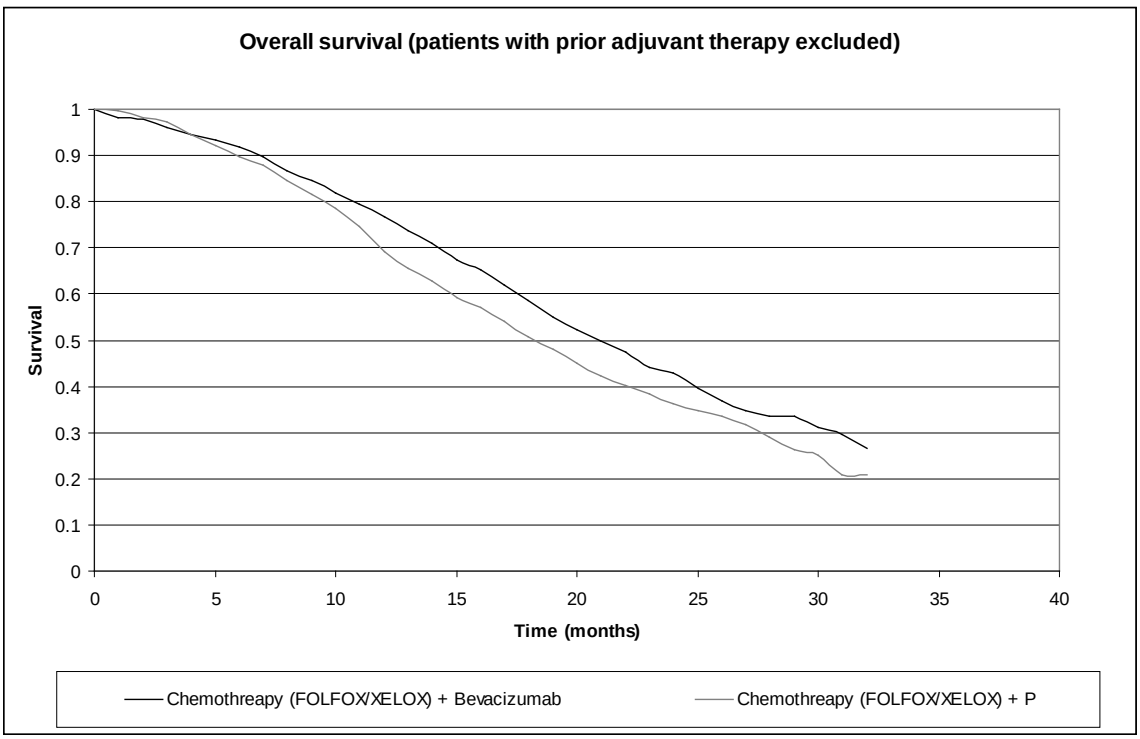
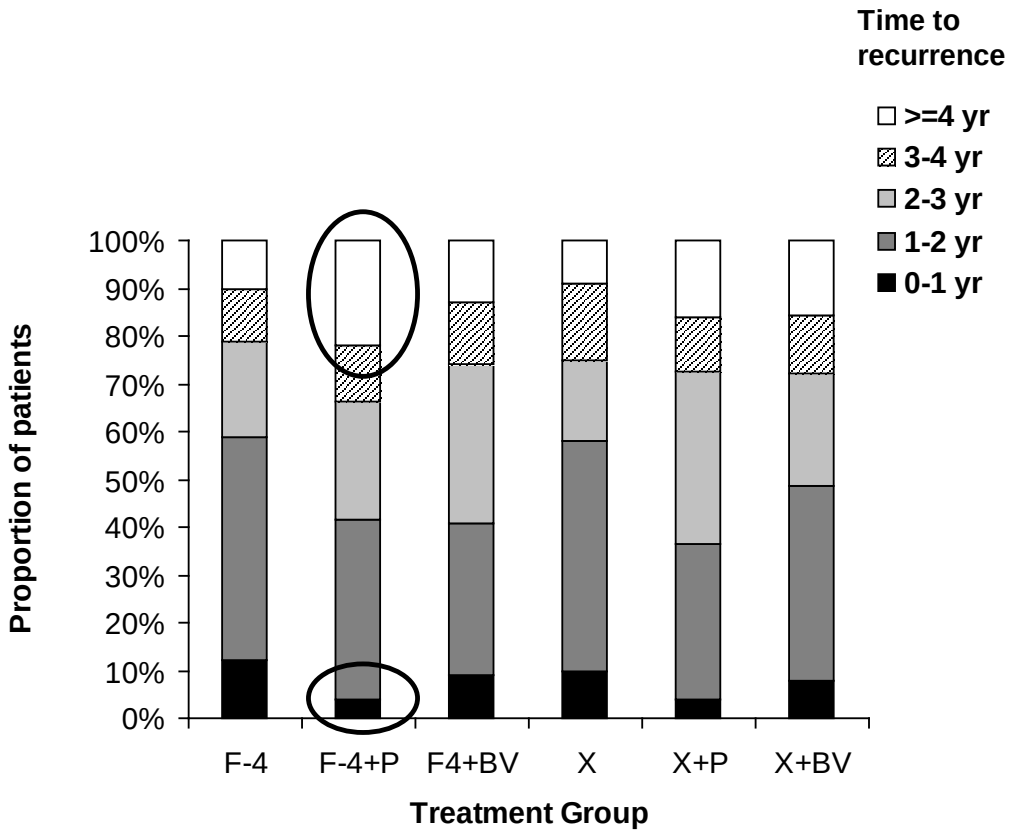


Figure 3 **Study NO16966: Distribution of patients according to time from adjuvant treatment to randomization by treatment arm**



Abbreviations: F-4, FOLFOX; F-4+P, P-FOLFOX; F4+BV, B-FOLFOX; X, XELOX; X+P, P-XELOX; X+BV, B-XELOX

5.3.2 Subgroup with liver metastases

The MS looks at the impact of adding bevacizumab to oxaliplatin-based chemotherapy in the subgroup of patients with liver metastases in trial NO16966. They comment that the addition of bevacizumab appears to increase both R0 resection rates and outcomes after resection. The manufacturers were requested, if possible, to perform a cost effectiveness analysis for this subgroup of patients but no analyses were provided.

5.3.3 Indirect/mixed treatment comparisons

The MS provides supportive evidence from an MTC undertaken by Golfinopoulos et al. 2007.[13] The MS supplementary data states that the MTC meta-analysis included results from the ECOG E3200 trial (second-line setting) and the pooled analysis of the 2x2 part of the NO16966 trial

(first-line setting). This MTC was not used as it combines data from first and second-line trials and head to head data regarding the interventions within the scope were available.

5.4 Extrapolation of clinical outcomes

The base-case in the MS used the Kaplan-Meier data for the first part of curve (up to median survival of 28 months) and then extrapolated beyond this point using a fitted parametric curve. In addition a sensitivity analysis was provided in which the entire time period was modelled using a parametric curve.

The MS truncated data at median follow up time (28 months). After median follow-up time (28 months) there were 14% (n=96) and 16% (n=211) patients alive in the XELOX/FOLFOX and XELOX/B-FOLFOX arms respectively. The ERG group requested the use of untruncated data to calculate Weibull parameter estimates as the method of fitting the parametric curves to survival data should allow for the greater uncertainty present in the tail of the curve. These were provided in the MS supplementary data.

The MS fits a Weibull distribution to the OS data. The MS describes three phases of the PFS curve (MS p119) but the ERG believed these to be somewhat subjective; whilst an exponential distribution rather than a Weibull distribution which was used for OS. The ERG suggested fitting a Weibull distribution to the PFS data from month six onwards and using this Weibull from month 28 onwards. The MS supplementary data took this approach but in fact used the Weibull from month six onwards.

The visual fit of the parametric curves to the Kaplan Meier curves is good as shown on p117 of the MS supplementary data.

Table 10. Method of extrapolation used for after median follow up (28 months)

	PFS	OS
MS	An Exponential with average hazard for months 13-28	A Weibull was fitted to data from months 1-28
MS supplementary data (analysis requested by ERG)	A Weibull was fitted to untruncated data from month 6 onwards	A Weibull was fitted to untruncated data

In the MS base-case a treatment effect is assumed to continue beyond median follow-up. A sensitivity analysis was performed which included no treatment effect after median follow up. This was achieved by applying the same risk of death in bevacizumab arms as in the chemotherapy alone arm beyond the point of median follow-up. This variable (no_treat_pf) does not seem to have been implemented in the model. (This variable does not change in the ICERs in the one-way sensitivity analysis.)

5.5 Health related quality of life

A utility value was assigned to each of the states of the Markov model. Patients receiving chemotherapy were assumed to have the same utility values regardless of treatment regimen.

The manufacturer undertook a systematic literature review in April 2007 to obtain utility values, which is described in the MS supplementary data Appendix B. The values found in the systematic review were not used and the manufacturer states that the reason for this was because they did not conform to the NICE reference case for utility values.

The original MS gives the source for the utility values used in the modelling as the cetuximab STA.[6] The ERG group requested the original source of utility values but this was not provided.

The MS includes a sensitivity analysis which varies utility values by +-20%. Given the uncertainty surrounding utility values the ERG considers this to be an appropriate analysis, although it is unlikely that the values will increase.

5.6 Resources and costs

Monthly costs for drug acquisition, pharmacy, drug administration, and monitoring were calculated as per treatment cycle costs multiplied by the average number of cycles per month as observed in NO16966. These monthly costs were then applied to the monthly model cycles to estimate the mean resource use per patient for each health state.

Kaplan-Meier survival analysis was used to calculate the mean treatment duration (i.e. time spent in the progression free survival on treatment (PFS_T) health state) based on the time from first dose to the time until cessation of treatment as recorded in the NO16966 trial.

Average adverse event and central venous access device (CVAD) costs per patient were applied to month one of the PFS_T health state and thus no discounting was applied.

The monthly drug acquisition cost of bevacizumab was only applied to the first 12 model cycles of the PFS_T health state to account for the 12 month price cap available through the Avastin Patient Access Scheme (APAS).

Given the very similar proportion of patients that received each of the post protocol treatments recorded in the NO16966 study (see MS appendix E5), it was assumed there were no differences in costs for second- and third-line treatments between the different intervention/comparators. Hence no cost for second- and third-line treatments has been applied in the model. Instead a monthly supportive care cost of £600 was applied for each of the interventions for the duration of post progression survival.

5.6.1 Dose received

For each drug the mean dose received in the trial is used to calculate relative dose intensity (RDI) observed in the trial which is used to calculate drug costs. To clarify, the dose table presented on p133 of the MS includes a bevacizumab dose for the XELOX arm but this is due to doses given in error.

5.6.2 Dose interruptions

Dose interruptions result in a longer cycle length and a smaller number of cycles administered per month. In the MS, cycle lengths for the XELOX and B-XELOX regimens have been pooled. The number of cycles administered per month is used to calculate treatment costs.

5.6.3 Treatment duration and continuation with bevacizumab after stopping oxaliplatin

Treatment duration (mean and median 6-7 months) in the trial was shorter than time until progression. Treatment duration varied by treatment arm, being longer with the addition of bevacizumab, and also longer on FOLFOX than on XELOX. The MS modelling is based on Kaplain Meier (KM) data (taken from trial) and XELOX and FOLFOX arms are modelled separately.

The ERGs clinical advisors suggested that in practice chemotherapy treatments would be likely to be stopped gradually rather than all at the same time. For example, oxaliplatin may be stopped

whilst other drugs are continued. However the MS comments that in the N016966 trial treatment with bevacizumab was often stopped at the same time point as the base chemotherapy was stopped. The manufacturers were asked to provide details of the number of patients for whom treatment with bevacizumab continued after chemotherapy was stopped but these data were not provided. The ERG assumes that this is possibly because very few patients continued.

The MS supplementary data provides treatment duration curves for each of the drugs included in a given regimen and there is some difference in the time-points that treatment with each drug in any one regimen was stopped. The mean treatment durations taken from the model are presented in Table 11. It is unclear whether the 2x2 data or all 6 arms have been used to calculate the treatment durations as the model uses the headings P-XELOX and P-FOLFOX but the figures (2&3) in the MS supplementary data are labelled XELOX+-P and FOLFOX+-P.

Table 11 Treatment durations with each component of the chemotherapy regimens

Regimen	Component	Mean treatment duration for component (area under KM curve), months
P-XELOX	oxaliplatin	
P-XELOX	capecitabine	
P-FOLFOX	oxaliplatin	
P-FOLFOX	5-FU	
B-XELOX	oxaliplatin	
B-XELOX	bevacizumab	
B-XELOX	capecitabine	
B-FOLFOX	oxaliplatin	
B-FOLFOX	bevacizumab	
B-FOLFOX	5-FU	

The treatment duration for each of the components of the regimen have been included in the model and the ERG believe the approach is reasonable. For simplicity the modelling assumes that oxaliplatin treatment duration is the same as bevacizumab treatment duration on the +bevacizumab arms. This will have little impact on the results as oxaliplatin is received free of charge on these arms. The model estimates the proportion of patients receiving for example capecitabine at a given time as the ratio of the Kaplan Meier capecitabine treatment duration to Kaplan Meier PFS multiplied by the estimate of PFS from the Weibull model. In the scenario in which XELOX and FOLFOX arms were pooled for survival modelling, un-pooled data was used to model treatment duration which is inconsistent.

5.6.4 Drug wastage

Drug wastage was not adjusted for in the original MS but has been included within the supplementary information. The APAS means that bevacizumab is a fixed price so adjusting for bevacizumab wastage is not necessary. Incremental costs will only be affected if oxaliplatin is associated with wastage as oxaliplatin is free on APAS and adjusting for oxaliplatin wastage and would slightly reduce the incremental cost of adding bevacizumab.

5.6.5 APAS

At the time of writing the approval/rejection of the APAS was still to be decided.
Bevacizumab is fixed price per patient £1200 per three week cycle/£800 per two week cycle, and free after year one. Oxaliplatin is free for patients receiving bevacizumab.

The MS states that the APAS will only be applicable for first-line metastatic CRC patients. If a patient has progressed (by the RECIST criteria (Solid evaluation criteria in solid tumours)) then the scheme would no longer apply as they would no longer be considered first-line.

The cost of administering the APAS scheme has been included in the modelling.

5.6.7 Adverse events

The incidence of adverse events used for the model was based on the occurrence of adverse events in the NO16966 trial. Our clinical advisors believe that the addition of bevacizumab is unlikely to reduce the incidence of adverse events. The ERG commented that the values provided in Table 35 showed a significant decrease in the incidence of several adverse events e.g. incidence of neutropenia/granulocytopenia is 44% with FOLFOX and 2% with B-FOLFOX. The values provided in Table 35 of the MS were incorrect and a new table was provided within the MS supplementary data.

In the Saltz et al.,[5] paper adverse events of special interest to bevacizumab with incidence greater than or equal to 2% were venous thromboembolic events, hypertension, bleeding, and arterial thromboembolic events (including ischemic cardiac events). Bleeding and arterial thromboembolic events had 2% incidence rate in some arms, but were not included within the modelling. The MS suggests that this is unlikely to have a significant effect on the ICER but the ERG cannot comment as the costs of treating bleeding and arterial thromboembolic events were neither provided nor reviewed.

The unit costs for adverse events are described with references (MS Table 34, p139). The ERG requested that details of the procedure/treatment/drugs which are included in these costs be

provided but none were given. It was therefore not possible to obtain clinical advice on the appropriateness of the procedure/treatment/drugs included for each adverse event.

5.7 Sensitivity analyses

5.7.1 One way sensitivity analyses

The MS includes an appropriate set of univariate sensitivity analyses and presents the results in tabular form and additionally as tornado diagrams.

5.7.2 Probabilistic sensitivity analysis (PSA) distributions

The distributions used for the PSA are described in MS Appendix E3.

Adverse event frequencies are modelled using a Beta distribution with parameters taken from trial data. A Gamma distribution was used for uncertainty in administration and monitoring costs. The standard error of the mean (SEM) should be used to estimate uncertainty which is calculated using N, the number of observations, but it is unclear whether the column entitled “se” in Table 53 describes the SEM.

The MS used a beta pert distribution to estimate uncertainty in adverse event costs. The motivation for using a beta pert is unclear particularly as a Gamma distribution was used for administration and monitoring costs. The method for calculating the beta distribution parameters was not described but a minimum and a maximum value chosen to be 50% and 150% of the mean cost were used. The ERG note that it would be better if the uncertainty was based on the data so, as it is a cohort model, the SEM should be used to describe uncertainty. The ERG requested clarification on the reasons for the choice of distribution used and the methods used for parameter calculation /selection but none were provided. The ERG believe that the distribution used is likely to overestimate the uncertainty in the adverse event costs.

A Beta distribution that fits to the confidence intervals of the utility value data was used to model the uncertainty in the utility values. Given the variation present between CRC utility values from the literature review the distributions used in the MS are not considered to adequately reflect the uncertainty in the utility values, which will result in the overall uncertainty being under estimated.

To represent uncertainty in the Kaplan Meier curves for PFS, OS and treatment duration a beta distribution was used for each of the transition probabilities. The transition probability at month n was sampled from a Beta (number failed in month n, number survived in month n) distribution.

The Kaplan Meier survival at month n was then calculated as the product of one minus the transition probability for each of the previous months. The uncertainty in the Weibull parameters was included using multivariate normal distributions.

5.7.3 Running PSA

The PSA macro is run using the Excel model drop down menu. The macro in the first version of the model did not initially work until a row reference error was corrected by the ERG. The macro performs 1000 runs of the PSA in a couple of minutes using a desk top with Intel Pentium 4 CPU 3.40 GHz, 2GB RAM. To evaluate whether 1000 configurations of parameter values was sufficient to provide robust results the ERG reran the analysis four times. The difference between the lowest and highest ICERs obtained from these runs was approximately £850.

5.7.4 PSA Results

The main results presented in the MS and in the executive summary are deterministic rather than the means from the probabilistic sensitivity analysis (PSA). The PSA results presented included the mean ICER, cost effectiveness plane, and cost effectiveness acceptability curve (CEAC).[30] At the request of the ERG the mean and 95% percentiles for the incremental costs, the incremental QALYs and the ICER were also presented but unfortunately these were only presented for the PSA results from the first version of the model (model 1, p49).

5.8 Model validation

The MS (p146) states that internal validation and debugging of the model was performed by an external company specialised in the development and validation of decision analytic models used for health economic analyses. The company had not been involved in the development of the model. The following validation procedures were performed:

- A check of completeness of reported results (health outcomes, economic outcomes) as compared with other published economic evaluations targeting the same indication; Tappenden et al, cetuximab for 1st line treatment of metastatic CRC[6].

- Execution of selected extreme tests to check the plausibility of model outcomes. Extreme testing was applied to the following parameters: treatment efficacy, adverse event costs, cost of study drugs and administration, discount rates, and health utilities.

- The methods of extrapolation (exponential and Weibull) were replicated and verified. Cost and utility inputs were validated with the evidence submission report.

To externally validate the model the estimated PFS and OS for the bevacizumab containing regimens estimated by the model were compared with real life data from the BRiTE and BEAT observational data. These data were also considered in light of the median results of the NO16966 study. The ERG note that the MS did not provide details of whether the model results and observational data compared were close.

5.9 Budget Impact

The MS includes a budget impact in which calculations are based on the assumption that B-XELOX is recommended for all first-line metastatic CRC patients suitable for this regimen and B-FOLFOX-6 is recommended in patients unsuitable for capecitabine. The current use of the different regimens is approximated using the results of a market research exercise which surveyed 38 oncologists who described the treatment of a total of 225 patients.

The budget impact assumes that only patients receiving FOLFOX, XELOX or FOLFIRI will be eligible for B-FOLFOX or B-XELOX. The ERG clinical advisors suggested that the group of eligible patients may in fact be larger and that patients who would currently receive single agent therapy with capecitabine or 5-fluorouracil may be offered capecitabine + bevacizumab or 5-fluorouracil plus bevacizumab if bevacizumab was available.

5.10 Results included in the MS

Table 12 summarises the ICER results presented in the MS and various sensitivity analyses are presented in Table 13. The main differences between the MS and the analysis provided with the MS supplementary data requested by the ERG are described in Section 5.1. The reason that these two sets of results are markedly different is due to more accurate modelling of treatment durations.

The costs and ICERs are greatly changed with the introduction of the APAS scheme. For example, the ICER for XELOX versus XELOX+B from the original analysis presented by the manufacturer changes from £82,098 to £34,170 on introduction of the APAS.

The results are sensitive to the trial data used to model survival (all 6 arms, or 2x2), whether or not XELOX and FOLFOX are pooled, and whether or not patients with prior adjuvant therapy are excluded. The ERG believe that the most appropriate base-case would use unpooled data from the 2x2 part of the N016966 trial with patient with prior adjuvant therapy excluded, however the manufacturer did not provide this analysis.

One way sensitivity analyses are presented in Table 13 and demonstrate the sensitivity of the ICERs to selected parameters and assumptions.

The ERG notes that the PSA was only run for one scenario. However, the ERG note that the mean ICER from the PSA is observed to be similar to the mean ICER from the deterministic analysis and thus the likely PSA results can be inferred from deterministic values. The ERG group requested that the 95% percentiles be presented with the PSA results. For FOLFOX the mean ICER is £41,518 and the 95% percentiles are (£31,136 £67,859). The percentiles around the incremental costs [redacted] demonstrate that there is considerable uncertainty around the expected QALYs. The ERG note that as the distribution used underestimate the uncertainty in utility values the uncertainty in the PSA will also be underestimated.

The exploratory second-line analysis performed by the manufacturer resulted in an ICER of £101,048 for B-FOLFOX4 versus FOLFOX4. In this analysis bevacizumab costs were £1600 per cycle and oxaliplatin was free with bevacizumab. The ERG understand that the APAS is only applicable to first-line therapy. This analysis seems to use the APAS but this is unclear as the cost of bevacizumab seems to be doubled at £1600 per cycle. The reasons for these assumptions are

unclear. The MS states that the far higher ICER's in the second-line setting to those of the first-line are most likely driven to a large extent by the higher dose which was used in the E3200 study relative to that used in the NO16966 study.

The budget impact provided in the MS assumes that B-XELOX is recommended for all first-line metastatic CRC patients suitable for this regimen and B-FOLFOX-6 is recommended in patients unsuitable for capecitabine. The MS estimates a budget impact of approximately £6 million in 2010 rising to £12 million in 2014.

Table 12 Results included in the MS (all results are with the APAS unless stated)

Scenario	Total costs(£s) and QALYs per patient	XELOX	FOLFOX-6	FOLFOX-4	FOLFIRI-mdG	FOLFIRI-dG	B-XELOX	B-FOLFOX-6	B-FOLFOX-4	ICI B-X XE
MS original analysis										
Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
PSA results for above analysis									95% percentiles of ICERs	
Above analysis without APAS [CIC]	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
MS supplementary data, requested by ERG										
Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
PSA results for above analysis									(95% percentiles not provided)	
Analysis using the 2x2 part of N016966, XELOX and FOLFOX arms pooled	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Analysis using the 2x2 part of N016966, XELOX and FOLFOX arms pooled, without prior adjuvant treatment	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Analysis using 2x2 part of N016966, XELOX and FOLFOX arms unpooled *	Costs(£s)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	

*The MS states that this analysis "uses truncated and oxaliplatin" - the ERG are unclear of the meaning of this.

Table 13 One-way sensitivity analysis of changes to mean parameter estimates

Analysis using data from all 6 arms of N016966, XELOX and FOLFOX arms pooled

				B-XELOX vs XELOX		B-FOLFOX-6 vs FOLFOX-6	
Parameter modified	**Base value**	**Low value**	**High value**	**ICER Low value**	**ICER High value**	**ICER Low value**	**ICER High value**
				base-case £35,912		base-case £ 36,569	
Utility Values							
PFS_T Utility value	0.77	0.616	0.924	£38,689	£33,507	£40,252	£33,505
PFS_{PT} Utility value	0.79	0.632	0.948	£39,274	£33,080	£39,145	£34,312
Progression Utility Value	0.68	0.544	0.816	£37,510	£34,444	£38,202	£35,071
Survival Analysis							
Weibull OS Survival curves (1) or mix of KM and Weibull (0)	0	0	1	£35,912	£38,802	£36,569	£39,542
Weibull PFS Survival curves (1) or mix of KM and Weibull (0)	0	0	1	£35,912	£34,526	£36,569	£33,668
assume treatment effect post follow-up 0 = yes 1 = no	0	0	1	£35,912	£35,912	£36,569	£36,569
Time horizon (years)	8	5	10	£39,768	£35,777	£40,537	£36,431
Clinical Practice Assumptions							
% pts requiring hospital transport	30%	0%	50%	£35,847	£35,955	£36,454	£36,646
% pts with CVAD insertion 0 = UK expert opinion, 1=recorded in trial	0	0	1	£35,912	£36,146	£36,569	£36,145
% FOLFOX pts with ambulatory pump	-	-	-	-	-	£36,569	£36,569
Unit Costs							
Cost of CVAD installation	£502	£301	£703	£35,911	£35,913	£36,511	£36,628
Cost of hospital funded transport per visit	£29	£18	£41	£35,886	£35,938	£36,523	£36,616
Cost per consultation with oncologist	£125	£75	£175	£35,469	£36,354	£36,088	£37,051
Cost of a CT scan	£135	£81	£189	£35,694	£36,129	£36,351	£36,788
Cost of administration day 1 of cycle	£317	£190	£444	£34,736	£37,087	£34,906	£38,233
Pharmacy cost (complex infusion)	£42	£25	£59	£34,875	£36,948	£34,555	£38,584
Pharmacy cost (simple infusion)	£25	£15	£35	£35,837	£35,986	£36,437	£36,702
Cost of Progressive Disease Health State	£600	£360	£840	£35,007	£36,817	£35,741	£37,398
Total bev, cape, ox Adverse Event costs	£248	£149	£347	£36,502	£35,321	-	-
Total XELOX Adverse Event costs	£334	£200	£467	£35,117	£36,706	-	-
Total FOLFOX Adverse Event costs	-	-	-	-	-	£35,365	£37,774
Cost of administration day 2 of cycle	-	-	-	-	-	£36,569	£36,569
Cost of inpatient stay of administration	-	-	-	-	-	£36,569	£36,569
Cost of 5-FU pump	-	-	-	-	-	£36,386	£36,753
Total B-FOLFOX Adverse Event costs	-	-	-	-	-	£37,541	£35,598
Cost of district nurse visit	-	-	-	-	-	£36,374	£36,765

5.11 Modelling Critique

5.11.1 Treatment duration

In the N016966 trial treatment was often stopped early and this may be a reason that a large difference in survival was not shown. The summary of product recommends that if oxaliplatin is stopped due to toxicity then bevacizumab and 5-fluorouracil should be continued until progression. The N016966 protocol specified that if one of the regimen components was discontinued due to toxicity, treatment could be continued with the remaining components.

Saltz et al.,[5] comments: "The reasons for the lack of treatment with bevacizumab or with chemotherapy until progression on this trial are not clear. One possibility is that when a cumulative toxicity, such as neurotoxicity or fatigue, reached a point at which the patient may have requested drug discontinuation, some investigators may not have fully appreciated that the protocol specifically permitted the discontinuation of one or more drugs while allowing for the continuation of others. Thus, for example, while discontinuation of oxaliplatin with continuation of fluoropyrimidine and bevacizumab was permitted, our analysis shows that this course of action was rarely taken."

The MS appears to accurately model treatment duration as it occurred in the trial but there is reason to believe that a longer duration of treatment with bevacizumab and 5-fluorouracil may be seen in clinical practice than was seen in the trial. Differences in treatment duration have a marked impact on the resulting ICERs.

5.11.2 Health related quality of life

The HRQoL literature review includes several studies such as Hamashima 2002[31] and van den Brink 2004[32] which calculate EuroQoL-5D (EQ-5D) from the societal perspective (see MS supplementary data Appendix B Table 2) . It is unclear to the ERG why these studies were considered by the manufacturer to not conform to the NICE reference case.[3] Other values presented include: van den Brink 0.11-0.9, Hamashima post operative rectal cancer 0.87+-0.22, Ness standard gamble Stage IV 0.24-0.27.

The source for the utility values used in the modelling is given as the cetuximab Single Technology Appraisal (STA)[6] in the MS. The ERG group requested the original source of utility values but this was not provided. The cetuximab STA describes several sources including unpublished data from the Crystal study and the Jonker[33] study, which is referenced for third line, but this is then stated to be incorrect in the clarification for cetuximab STA. A further source

mentioned is an abstract (Mittmann)[34] which doesn't include the HRQoL data. The response to the cetuximab STA[35] clarification provides a poster which was not available to the ERG which contained Health Utility Index (HUI) data up to week 24, (the data used claimed to be taken from an average of values up to week 40). It is unclear whether the HRQoL data used is HUI or EQ-5D. On p115 Bidard et al 2008[36] is referenced but there is no mention of quality of life in the abstract. Therefore the ERG were unable to adequately check the sources of the utility values, due to this poor referencing and commentary.

To summarise, the sources of the utility values used in the model have been poorly referenced so the ERG have been unable to adequately check them. The MS does not clearly demonstrate how the utility values used were selected and why other utility values identified during the literature review were not used. Given the variation present between CRC utility values from the literature review the distributions used are not considered to adequately reflect the uncertainty in the utility values.

The MS assumes that the utility value when on treatment is 0.77 regardless of whether or not the patient is receiving bevacizumab. As the bevacizumab arms of the N016966 trial were associated with more adverse events (e.g. hypertension, thromboembolic events, see Section 4.2.1) this may be unfairly favourable to the bevacizumab arms.

Without reference to the absolute utilities, our clinical advisors commented that the utility values from the cetuximab trial could be relevant to patients receiving bevacizumab. When treatment was stopped the MS assumes the same utility value as the general population for the time until progression. However, our clinical advisors additionally believe that the assumption that patients who have finished treatment but have not progressed have the same utility as the general population is unrealistic, as such patients are often mentally and physically less fit after six months of chemotherapy than the general population. The MS assumes HRQoL is higher off treatment pre-progression due to cessation of adverse events. Accordingly the ERG comment that the value of 0.77 for patients on treatment may be an overestimation as the general population value for this age group is 0.79.

The ERGs clinical advisors also commented that the XELOX regimen is considered to be more convenient than the FOLFOX regimen and this may affect HRQoL and this has not been included in the modelling. Disutility of adverse events was not explicitly included within the modelling but it is assumed that the on treatment utility values implicitly include any disutility due to adverse events. The ERG notes that the utility value for progressive disease seems quite high; it is

possible that a lower utility value be seen in the last few weeks of life. The ERG notes that the MS did not use age specific utility values.

5.11.3 Efficacy

The MS presents several scenarios using the different efficacy results i.e. all 6 arms, 2x2 only, pooled and un-pooled XELOX and FOLFOX, and excluding patients with prior adjuvant therapy. These different scenarios result in markedly different ICERs, see Table 12.

The MS presents several scenarios using different methods of modelling and extrapolating the survival curves. These scenarios result in markedly different ICERs, see Table 12.

5.11.6 Intermittent versus continuous chemotherapy

In England chemotherapy for metastatic colorectal cancer is currently administered either "continuously" or "intermittently". During an "intermittent" treatment programme chemotherapy may be administered for say 3-6 months and then stopped and restarted at progression. (It is unclear whether this further treatment at progression is termed first-line or second-line). The MS states that continuation of treatment would still be regarded as first-line if the patient has not progressed (by RECIST criteria).

One of the ERGs clinical advisors suggested that if bevacizumab were added to existing chemotherapy then they would use a continuous treatment programme as this is in line with the evidence base for bevacizumab. The other clinical advisor agreed that a continuous treatment programme would more often be used in this case. The MS also suggests that that treatment with bevacizumab should be continuous.

In the model continuous treatment with and without bevacizumab are compared (as used in the trial). In clinical practice it may be the case that intermittent treatment is replaced by continuous treatment plus bevacizumab. It is not clear what the differences will be (in terms of costs and QALYs) between continuous and intermittent treatment.

5.11.7 Cycle lengths

The N016966 trial showed several differences which occurred with the addition of bevacizumab. The duration of treatment was longer with the addition of bevacizumab. Cycle lengths are slightly shorter with bevacizumab and the confidence intervals presented in MS supplementary data Table 2 suggest this difference could be significant. If cycle lengths were modelled separately for all arms then this would reduce cycle lengths for comparator arms and increase cycle length for

bevacizumab arms resulting in a higher monthly cost for the addition of bevacizumab. For example the number of cycles per month would be 1.81 for P-FOLFOX and 1.87 for B-FOLFOX.

5.11.8 Model accuracy

Several differences were found when comparing model inputs to values in the MS Table 23 and are described in Table 14. The manufacturer has confirmed that the values used in the model are the correct values.

Table 14 Discrepancies between MS and model inputs

Parameter	Value in report	Value in model
Progressive disease	1-(PFS+OS)	OS-PFS
Drug admin and pharmacy costs (different for all regimens) e.g. B-FOLFOX4 inpatient	£988 per cycle	£1363 per cycle
Cycles per month 5FU	1.83	1.84
Utility value post progression	0.67	0.68
Chemo+bev OS Lambda	0.0059	0.0046
Chemo+bev OS Gamma	1.5473	1.6357
Chemo+bev PFS Lambda	0.0248	0.0238
Chemo+bev PFS Gamma	1.4584	1.4878
Chemo OS Lambda	0.0070	0.0052
Chemo OS Gamma	1.5473	1.6357
Chemo PFS Lambda	0.0311	0.0284
Chemo PFS Gamma	1.4584	1.4878

5.12 Summary of uncertainties and issues

The main issues and areas of uncertainty identified by the ERG are listed here and are discussed in more detail within the model critique and additional analyses sections.

5.12.1 Efficacy data

The restriction to trial data from the 2x2 part, the pooling of the XELOX and FOLFOX arms, and the exclusion of patients with prior adjuvant therapy from the data set all have a very large impact on resulting ICERs.

It is unclear whether the analysis excluding prior adjuvant patients and the resulting ICERS are relevant for the whole patient group or should be restricted to the subgroup who have not received prior adjuvant therapy. If survival differences were solely related to differences in time since prior adjuvant therapy then performing this analysis should remove the remove the bias this caused while still being relevant to the entire patient group. If the survival difference is due to a difference in the efficacy of bevacizumab in the prior adjuvant therapy patient group then the resulting ICER would just apply to the subgroup of patients who have not received prior adjuvant therapy.

The ERG would recommend an analysis in which 2x2, unpooled, excluding prior adjuvant patients data is used. This was not provided in the MS and data was not available for the ERG to perform this analysis.

The different approaches for modelling the survival data (Kaplan Meier and fitting a Weibull distribution) both result in slightly different ICERs. As the Weibull distributions have such a good visual fit to the Kaplan Meier curves the change to the ICERs is generally small.

5.12.2 HRQoL data

The MS does not make use of the range of utility values identified from the literature review and does not explain why these values were not used. The sources of the utility values used in the MS were poorly referenced resulting in the ERG being unable to check them. There is also uncertainty around the clinical plausibility of the post treatment pre-progression utility value. The distributions used for the utility values in the PSA reflect the uncertainty relating to the specific values used but underestimate the uncertainty relating to the selection of utility values. Using wider distributions for utility values would significantly increase the confidence intervals around the mean ICERs from the PSA.

5.12.3 Treatment duration

The ERGs clinical advisors suggest that in clinical practice treatment with non-oxaliplatin chemotherapy components may continue beyond oxaliplatin cessation unlike in the N016966 trial where this rarely happened. Due to the structure of the APAS scheme (in which oxaliplatin is received free of charge) this could have a marked impact on the ICERs. The ERG ran an exploratory analysis in which oxaliplatin was stopped one month before the other chemotherapy components.

5.12.4 Intermittent versus continuous chemotherapy

Our clinical advisors suggest that current care in England is often intermittent treatment with chemotherapy. The trial and the model both represent continuous treatment chemotherapy. The difference in cost and effectiveness between intermittent and continuous treatment is unknown so it is unclear how this difference could affect the ICERs. Literature suggests that intermittent treatment has a similar efficacy to continuous treatment and may be associated with better HRQoL.[37] It is possible that intermittent treatment may be cheaper than continuous treatment whilst having a similar efficacy. If this was the case then the ICER for continuous treatment with bevacizumab versus intermittent treatment would be greater than the ICER for continuous treatment with bevacizumab versus continuous treatment presented here. The ERG recommends that this is an area for further research.

5.12.5 Cycle lengths

The MS pooled with and without bevacizumab treatment arms when calculating cycle lengths. The data suggests significantly shorter cycle lengths with the addition of bevacizumab so it may not be appropriate to pool the data here. The ERG ran additional analysis in which cycle lengths were not pooled across trial arms.

6. Additional work undertaken by the ERG

We recall that the main conclusions from the analyses presented in the MS were:

- Running the analysis without the APAS resulted in much larger ICERs for adding bevaicuzumab.
- Restriction to data from the 2x2 part of the trial, pooling of XELOX and FOLFOX arms, and exclusion of patients with prior adjuvant therapy, all had a significant impact on the ICERs for adding bevacizumab.

The results of additional analyses undertaken by the ERG are provided in Tables 15 and 16. The main conclusions are that:

- Decreasing all three utility values by 20% had a significant impact on the ICERs.
- Reducing drug costs to correspond with stopping oxaliplatin one month earlier in all arms significantly increased the ICERs.

6.1 Extrapolation of survival curves

An examination of the survival curves shows that for each of the comparisons B-XELOX vs. XELOX and B-FOLFOX vs. FOLFOX the curves for both OS and PFS appear to cross in the first 6 months. It may therefore be reasonable to use the Kaplan Meier curve up to this point and then use the fitted Weibull curve for the period after six months. The ERG ran an analysis in which the Kaplan Meier curve is used initially and the Weibull curve is used after month six. This did not have a significant impact on the ICERs.

6.2 HRQoL sensitivity analysis

There is likely to be a correlation between the utility values used in the model. There is likely to be a constant ranking, for example: ‘utility pre-progression off treatment’ would be greater than ‘utility pre-progression on treatment’, which would be greater than ‘utility post progression’, which may not be maintained in the univariate sensitivity analysis. The ERG group have run a sensitivity analysis in which all three utility values were reduced by 20%. Decreasing all three utility values by 20% caused a marked increase in the ICERs.

6.3 Exploratory analysis demonstrating effect of treatment duration on ICER

The ERGs clinical advisors suggest that in clinical practice treatment with non-oxaliplatin chemotherapy components may continue beyond oxaliplatin cessation unlike in the N016966 trial where this rarely happened. The ERG performed an exploratory analysis to investigate the impact this would have on the ICER.

In the exploratory analysis oxaliplatin was not administered in the last month of treatment (but all other drugs were continued: capecitabine, bevacizumab, 5FU etc) for all arms. This reduced costs by £[redacted] in the XELOX arm and by £[redacted] in the FOLFOX-6 arm. The costs relating to the B-XELOX and B-FOLFOX arms would remain constant as oxaliplatin is free for these arms. There may also be associated survival differences but these could be present in both arms; in this analysis there is assumed to be no change in incremental survival. This analysis greatly increases the ICER. The ERG note that the reason for the difference is related to the structure of the APAS scheme (i.e. that oxaliplatin is given for free but this benefit is removed if oxaliplatin is stopped earlier than bevacizumab). The ERG point out that if the difference in treatment duration was greater than one month the effect on the ICER would be even more pronounced.

6.4 Cycle length

Cycle lengths are slightly shorter with bevacizumab and the confidence intervals presented in the MS supplementary data Table 2 suggest this difference could be significant. If cycle lengths were modelled separately for all arms then this would reduce cycle lengths for comparator arms and increase cycle length for bevacizumab arms resulting in a higher monthly cost for the addition of bevacizumab. For example, the number of cycles per month would be 1.81 for P-FOLFOX and 1.87 for B-FOLFOX. The ERG ran an analysis in which the cycle durations were calculated separately for each treatment arm rather than pooled. This analysis resulted in higher ICERs for the B-FOLFOX vs FOLFOX comparison.

Table 15 Table of results of additional work undertaken by the ERG

All analyses use data from 2x2 part of N016966, XELOX and FOLFOX arms pooled, patients with prior adjuvant therapy excluded, with APAS				ICERs	
	Progression Free Survival (PFS) modelling	Overall Survival (OS) modelling	Scenario	B-XELOX vs. XELOX	B-FOLFOX6 vs FOLFOX6
MS	KM up to month 6 then Weibull	Weibull		£ 36,006	£ 31,174
ERG analysis	**KM up to month 6 then Weibull**	**KM up to month 28 then Weibull**	*****	**£ 36,354**	**£ 31,452**
ERG analysis	Weibull	Weibull		£ 35,135	£ 28,976
ERG analysis	KM up to month 6 then Weibull	KM up to month 6 then Weibull	As treatment arms stop crossing at month 6 this may be an appropriate point at which to start extrapolation	£ 36,438	£ 31,523
ERG analysis	KM up to month 6 then Weibull	KM up to month 28 then Weibull	-20% all utility values (0.63,0.62,0.54)	£ 45,443	£ 39,315
ERG analysis	KM up to month 6 then Weibull	KM up to month 28 then Weibull	Treatment duration - Oxaliplatin stopped one month earlier (assumed no change in incremental survival)	£ 43,511	£ 39,478
ERG analysis	KM up to month 6 then Weibull	KM up to month 28 then Weibull	Cycle lengths unpooled	£ 36,488	£ 32,900

*The ERG suggest that the results using the 2x2 part of the N016966 trial with XELOX and FOLFOX unpooled and prior adjuvant patients excluded may be the most appropriate base-case. As this analysis was not presented in the MS and data was not available for the ERG to perform it the analysis in bold may be the most appropriate analysis available. The one way sensitivity presented in Table 16 was run for this analysis.

Table 16 One-way sensitivity analysis of changes to mean parameter estimates

Base-case: 2x2 part of N016966 trial, XELOX and FOLFOX arms pooled, survival data modelled using KM then Weibull.

				B-XELOX vs XELOX		B-FOLFOX-6 vs FOLFOX-6	
Parameter modified	**Base value**	**Low value**	**High value**	**ICER Low value**	**ICER High value**	**ICER Low value**	**ICER High value**
				base-case	£ 36,354	base-case	£ 31,452
Utility Values							
PFST Utility value	0.77	0.616	0.924	£36,944	£35,783	£33,514	£29,630
PFSPT Utility value	0.79	0.632	0.948	£37,671	£34,832	£36,489	£27,637
Progression Utility Value	0.68	0.544	0.816	£36,359	£36,349	£31,467	£31,438
Survival Analysis							
Wiebull OS Suvival curves (1) or mix of KM and Weibull (0)	0	0	1	£36,354	£36,006	£31,452	£31,174
Wiebull PFS Suvival curves (1) or mix of KM and Weibull (0)	0	0	1	£36,354	£35,481	£31,452	£29,232
assume treatment effect post followup 0 = yes 1 = no	0	0	1	£36,354	£36,354	£31,452	£31,452
Time horrizon (years)	8	5	10	£39,599	£36,285	£34,082	£31,396
Clinical Practice Assumptions							
% pts requiring hospital transport	30%	0%	50%	£36,299	£36,391	£31,375	£31,504
% FOLFOX pts with ambulatory pump	0.95	0.5	1	£36,354	£36,354	£31,452	£31,452
% pts with CVAD insertion 0 = UK expert opion, 1=recorded in trial	0	0	1	£36,354	£36,628	£31,452	£30,965
Unit Costs							
Cost of CVAD installation	£502	£301	£703	£36,353	£36,355	£31,384	£31,521
Cost of hospital funded transport per visit	£29	£18	£41	£36,332	£36,376	£31,421	£31,483
Cost of 5-FU pump	£35	£21	£49	£36,354	£36,354	£31,329	£31,576
Cost per consultation with oncologist	£125	£75	£175	£35,852	£36,856	£30,950	£31,955
Cost of a CT scan	£135	£81	£189	£36,079	£36,629	£31,177	£31,727
Cost of district nurse visit	£37	£22	£52	£36,354	£36,354	£31,321	£31,584
Cost of administration day 1 of cycle	£317	£190	£444	£35,253	£37,455	£30,334	£32,571
Cost of administration day 2 of cycle	£227	£136	£318	£36,354	£36,354	£31,452	£31,452
Cost of inpatient stay of administration	£1,052	£631	£1,473	£36,354	£36,354	£31,452	£31,452
Pharmacy cost (complex infusion)	£42	£25	£59	£35,173	£37,535	£29,396	£33,509
Pharmacy cost (simple infusion)	£25	£15	£35	£36,291	£36,418	£31,363	£31,542
Cost of Progressive Disease Health State	£600	£360	£840	£36,344	£36,364	£31,538	£31,367
Total B Cape Ox Adverse Event costs	£248	£149	£347	£37,045	£35,664		
Total FOLFOX Adverse Event costs	£334	£200	£467	£35,425	£37,283		
Total B Cape Ox Adverse Event costs	£407	£244	£569			£32,585	£30,320
Total FOLFOX Adverse Event costs	£504	£303	£706			£30,048	£32,857

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The manufacturer's submission to NICE includes a systematic review of the clinical-effectiveness literature. Although two randomised controlled trials were identified, one as first-line therapy (patients not previously treated for their metastatic disease, NO16966 trial)[5] and one in second-line therapy (for previously treated patients with metastatic disease, E3200 trial),[4] the manufacturer is seeking approval for first-line use only (and therefore forms the main pivotal evidence in the submission). The manufacturer claims that they could not demonstrate a cost-effectiveness case for the use of bevacizumab in second-line therapy. The manufacturers' primary pooled analysis of superiority (using the intention to treat population) in the NO16966 trial showed that after a median follow up of 28 months, the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) significantly improved progression free survival and overall survival compared with chemotherapy alone (P-XELOX/ P-FOLFOX-4/ XELOX/ FOLFOX-4 combined) in adult patients with histologically confirmed metastatic colorectal cancer not previously treated (median progression free survival: 9.4 versus 7.7 months [absolute difference, 1.7 months]; hazard ratio, 0.79; p=0.0001; median overall survival: 21.2 versus 18.9 months [absolute difference, 2.3 months]; hazard ratio, 0.83; p=0.0023). The manufacturer's pooled analysis of non-inferiority (using the eligible patient population and the intention to treat population) showed that the XELOX (XELOX/P-XELOX/B-XELOX combined) and FOLFOX-4 (FOLFOX-4/ P-FOLFOX-4/ B-FOLFOX-4 combined) based regimens were equivalent for both progression free survival (p=not significant, values not reported) and overall survival (p=not significant, values not reported). Given the trial design limitations (two part study, open label design, imbalance of known prognostic factor [time between primary treatment and recurrence] and relatively short duration of chemotherapy treatment [aproximately 6 months] despite the fact that the trial protocol allowed coninuation of the study therapy until progressive disease or unacceptable toxicity) and the manufacturers primary pooled analysis of superiority (which pooled all patients in the trial), the ERG considers the manufacturers secondary pooled analysis of superiority (restricted to patients in the second 2x2 part of the NO16966 study, as per the original statistical trial plan and reported in the primary published peer reviewed clinical paper)[5] to be more appropriate. This analysis also showed that the addition of bevacizumab to chemotherapy (B-XELOX / B-FOLFOX-4 combined) improved progression free survival and overall survival compared with chemotherapy (P-XELOX/ P-FOLFOX-4 combined) alone (median progression free survival: 9.4 versus 8.0 months [absolute difference, 1.4 months]; hazard ratio, 0.83; p=0.0023; median overall survival: 21.3 versus 19.9 months [absolute difference, 1.4 months]; hazard ratio, 0.89; p=0.0769). No corresponding data

were provided or undertaken for the analysis of non-inferiority. A pre-defined subgroup analysis on progression free survival found that the statistical superiority of bevacizumab plus chemotherapy was evident in the XELOX subgroups (B-XELOX versus P-XELOX; hazard ratio, 0.80; 97.5% CI: 0.66 to 0.96; p=not reported) but did not reach the significance level in the FOLFOX-4 subgroups (B-FOLFOX-4 versus P-FOLFOX-4; hazard ratio, 0.89; 97.5% CI: 0.74 to 1.06; p=not reported). Additional post hoc exploratory analyses (following the results from the Adjuvant Colon Cancer End Points [ACCENT] study, which found that there was a significant and direct correlation between time to recurrence after surgery and survival after recurrence in patients whose disease recurred after surgery and adjuvant treatment) showed that removing the subgroup of patients that may have slower tumour progression after adjuvant treatment (an imbalance between treatment groups with regard to an important prognostic factor which was not recognised at the start of the NO16966 trial), significantly improved the hazard ratios for adding bevacizumab to chemotherapy compared with chemotherapy alone for both overall survival and progression free survival. Depending on the analyses conducted (e.g. exclusion of patients with adjuvant chemotherapy from all four treatment arms of the factorial study, or from FOLFOX groups only or from P-FOLFOX group only) the hazard ratios for overall survival ranged from 0.83 to 0.85 ($p<0.03$) and the hazard ratios for progression free survival ranged from 0.74 to 0.77 ($p<0.0001$). Although this may be plausible, the ERG note that caution should be exercised as this is a post hoc exploratory analysis.

7.2 Summary of cost effectiveness issues

A number of issues were identified that had an impact on the ICERs. These included the following:

- Avastin Patient Access Scheme(APAS): At the time of writing the decision on whether the proposed APAS scheme would be accepted was unknown. The majority of the analysis presented by the manufacturer included the APAS. Running the model without the APAS resulted in much higher ICERs.

- Efficacy data: It is unclear whether the clinical evidence from the RCT trial used in the MS should be pooled (without weighting for uncertainty) according to data from the initial two arm part and the 2x2 factorial part of the NO16966 study or restricted to patients in the 2x2 factorial part, as per the original statistical trial plan of the NO16966 trial. Additionally it is unclear whether patients with prior adjuvant chemotherapy should be excluded from the analysis. The restriction to trial data from the 2x2 part of the NO16966 trial, the unpooling of the XELOX and FOLFOX arms, and the restriction to the data of patients without prior adjuvant chemotherapy, all have a large impact on resulting ICERs. Restriction to the 2x2 part of the N016966 increased ICERs, exclusion

of patients with prior adjuvant chemotherapy decreased ICERs, and pooling the XELOX and FOLFOX arms affected the XELOX and FOLFOX ICERs in different directions.

- HRQoL data: The MS does not make use of the range of utility values identified from the literature review and do not explain why these values were not used. The sources of the utility values used in the MS were poorly referenced resulting in the ERG being unable to check them. There is also uncertainty around the clinical plausibility of the post treatment pre-progression utility value. The distributions used for the utility values in the probabilistic sensitivity analyses (PSA) reflect the uncertainty relating to the specific values used but underestimate the uncertainty relating to the selection of utility values. Using wider distributions for utility values would significantly increase the confidence intervals around the mean ICERs from the PSA. Reducing the utility values by 20% significantly increased the ICERs.

- Treatment duration: In clinical practice treatment with non-oxaliplatin chemotherapy components may continue beyond oxaliplatin cessation, however in the N016966 trial was rarely seen. Due to the structure of the Avastin Patient Access Scheme (APAS) (in which oxaliplatin is received free of charge) this could have a significant impact on the ICERs. The ERG ran an exploratory analysis to determine the effect on the ICER of stopping oxaliplatin only one month earlier and assuming incremental effectiveness is unchanged. This exploratory analysis significantly increased the ICERs.

- Intermittent versus continuous chemotherapy: The ERGs clinical advisors suggested that current care in England is often intermittent treatment with chemotherapy. The trial and the model both represent continuous treatment chemotherapy. The difference in cost and effectiveness between intermittent and continuous treatment is unclear. As an example, if intermittent treatment was cheaper than continuous treatment whilst having a similar efficacy, then the ICER for continuous treatment with bevacizumab versus intermittent treatment would be greater than the ICERs for continuous treatment with bevacizumab versus continuous treatment presented here.

7.3 Implications for research

- The N016966 trial protocol and the summary of product characteristics both allow continued treatment with bevacizumab after oxaliplatin cessation until disease progression. In clinical practice treatment with non-oxaliplatin chemotherapy components may continue beyond oxaliplatin cessation unlike in the N016966 trial where this rarely happened. Further research is

required to evaluate the likely duration of bevacizumab treatment in clinical practice and the survival associated with longer treatment duration. There are ongoing trials to evaluate the use of bevacizumab after cessation of chemotherapy in the metastatic phase DREAM-OPTIMOX3 trial[38], and as adjuvant therapy in the QUASAR2 study.[39]

- As the current care for a large proportion of patients receiving oxaliplatin-based regimens is "intermittent treatment", further research is required to evaluate the effect of changing these patients to "continuous treatment" with the addition of bevacizumab.

- Further research is required in finding ways to select patients who will benefit form bevacizumab therapy (e.g. this is analogous to the use of the Kirsten Rat Sarcoma [KRAS] test to select patients for cetuximab or the Human Epidermal growth factor Receptor 2 [HER2] for herceptin)

8. REFERENCES

1. Kabbinavar, F., Hurwitz, H. I., Fehrenbacher, L., Meropol, N. J., Novotny, W., Lieberman, G., Griffing, S., and Bergsland, Emily K. Phase II, randomized trial comparing bevacizumab plus fluorouracil (FU)/leucovorin (LV) with FU/LV alone in patients with metastatic colorectal cancer. *Journal of Clinical Oncology* 2003; **21** 60-65.

2. National Institute for Health and Clinical Excellence (NICE) Technology Appraisal Guidance 93 - Irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer. London, NICE 2005

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

4. Giantonio, Bruce J. Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: results from the Eastern Cooperative Oncology Group study E3200. *Journal of Clinical Oncology* 2007; **25** 1539-1544.

5. Saltz, L. B. C. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: A randomized phase III study. *Journal of Clinical Oncology* 2008; **26** 2013-2019

6. National Institute for Health and Clinical Excellence *Final Appraisal Determination: Cetuximab for first-line treatment of metastatic colorectal cancer.* Available at http://guidance.nice.org.uk/TA/Wave15/82 Accessed 02/09/09

7. Brettle, A. J., Long, A. F., Grant, M. J., and Greenhalgh, J. Searching for information on outcomes: do you need to be comprehensive? *Quality in Health Care* 1998; **7** 163-169

8. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

9. Hochster, H. S., Hart, L. L., Ramanathan, R. K., Childs, B. H., Hainsworth, J. D., Cohn, A. L., Wong, L., Fehrenbacher, L., Abubakr, Y., Saif, M. W., Schwartzberg, L., and Hedrick, E. Safety and efficacy of oxaliplatin and fluoropyrimidine regimens with or without bevacizumab as first-line treatment of metastatic colorectal cancer: results of the TREE Study. *Journal of Clinical Oncology* 2008; **21** 3523-3529.

10. Berry, S., Cunningham, D., and Michael, M. Preliminary efficacy of bevacizumab with first-line FOLFOX, XELOX, FOLFIRI, and fluoropyrimidines in mCRC: First BEAT. *Proc Gastrointestinal Cancers Symposium* 2008; January 25-27, 2008 (abst 350)

11. Kozloff, M., Sugrue, M., and Berlin, J. Safety and effectiveness of bevacizumab and chemotherapy in elderly patients with metastatic colorectal cancer: Results from the BRiTE prospective cohort study. *Proc Gastrointestinal Cancers Symposium, January 25-27, 2008 (abst 345)* 2008;

12. Cao, Y., Tan, A., Gao, F., Liu, L., Liao, C., and Mo, Z. A meta-analysis of randomized controlled trials comparing chemotherapy plus bevacizumab with chemotherapy alone in metastatic colorectal cancer. *Int J Colorectal Dis* 2009; **24** 677-685.

13. Golfinopoulos, V Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. *Lancet Oncology* 2007; **8** 898-901.

14. National Institute for Health and Clinical Excellence (NICE) Single Technology Appraisal (STA): Specification for manufacturer/sponsor submission of evidence - Update to reflect the new Guide to the Methods of Technology Appraisals. London, NICE 2008

15. Khan, K. S., Kunz, R., Kleijnen, J., and Antes, G. Systematic reviews to support evidence-based medicine – How to review and apply findings of healthcare research. *Royal Society of Medicine Press Ltd* 2003;

16. Schulz, K. F., Chalmers, I., Hayes, R. J., and Altman, D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273** 408-412.

17. Juni, P., Witschi, A., Bloch, R., and Egger, M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; **282** 1054-1060.

18. O'Connell, M. J., Campbell, M. E., Goldberg, R. M., Grothey, A., Seitz, J. F., Benedetti, J. K., Andre, T., Haller, D. G., and Sargent, D. J. Survival following recurrence in stage II and III colon cancer: findings from the ACCENT data set. *Journal of Clinical Oncology* 2008; **10** 2336-2341.

19. Juni, P., Altman, D., and Egger, M. Assessing the quality of controlled clinical trials. *British Medical Journal* 2001; **323** 42-43.

20. Hurwitz, H., Fehrenbacher, L., Novotny, W., Cartwright, T., Hainsworth, J., Heim, W., Berlin, J., Baron, A., Griffing, S., Holmgren, E., Ferrara, N., Fyfe, G., Rogers, B., Ross, R., and Kabbinavar, F. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England Journal of Medicine* 3-6-2004; **350** 2335-2342.

21. Kabbinavar, F. F., Schulz, J., McCleod, M., Patel, T., Hamm, J. T., Hecht, J. R., Mass, R., Perrou, B., Nelson, B., and Novotny, W. F. Addition of bevacizumab to bolus fluorouracil and leucovorin in first-line metastatic colorectal cancer: results of a randomized phase II trial.[see comment]. *Journal of Clinical Oncology* 2005; **23** 3697-3705.

22. Egger, M., Smith, G. D., and Phillips, A. N. Meta-analysis: principles and procedures. *British Medical Journal* 1997; **315** 1533-1537.

23. Lievre, M., Cucherat, M., and Leizorovicz, A Pooling, meta-analysis, and the evaluation of drug safety. *Current Controlled Trials in Cardiovascular Medicine* 2002; **3**

24. Bravata, D. M. and Olkin, I. Simple pooling versus combining in meta-analysis. *Eval Health Prof* 2009; **24** 218-230.

25. European Medicines Agency (EMEA). Assessment Report for Avastin. Procedure Number. EMEA/H/C/000582/II/0014. Available at *http://www.emea.europa.eu/humandocs/PDFs/EPAR/avastin/Avastin-H-582-II-14-AR.pdf* Accessed 6 September 2009

26. Cassidy, J., Clarke, S., Diaz-Rubio, E., Scheithauer, W., Figer, A., Koski, S., Lichinitser, M., Yang, T. S., and Rivera, F. Randomized phase III study of capecitabine plus oxaliplatin compared with fluorouracil/folinic acid plus oxaliplatin as first-line therapy for metastatic colorectal cancer. *Journal of Clinical Oncology* 2008; **26** 2006-2012.

27. Piaggio, G., Elbourne, D. R., Altman, D. G., Pocock, S. J., and Evans, S. J. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006; **295** 1152-1160.

28. Wagner, A. D., Arnold, D., Grothey, A. A., Haerting, J., Unverzagt, S., Wagner, Anna Dorothea, Arnold, Dirk, Grothey, Axel A. G., Haerting, Johannes, and Unverzagt, Susanne Anti-angiogenic therapies for metastatic colorectal cancer. *Cochrane Database of Systematic Reviews* 2009;CD005392-

29. Montgomery, A. A., Peters, T. J., and Little, P. Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol* 2003; **3** 26-

30. Fenwick, E, Claxton K, and Sculpher M Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* 2001; **10** 779-787.

31. Hamashima, C Long-term quality of life of postoperative rectal cancer patients. *Journal of Gastroenterology and Hepatology* 2002; **17** 571-576.

32. van der Brink, M., van der Hout, W. B., and Stiggelbout, A. M. Cost-utility analysis of preoperative radiotherapy in patients with renal cancer undergoing total mesorectal excision: a study of the Dutch colorectal cancer group. *J Clin Oncol* 2004; **22** 244-253.

33. Jonker, D. J. Cetuximab for the treatment of colorectal cancer. *New Engl J Med* 2007; **357** 2040-2048.

34. Mittmann, N., Au, H. J., and Tu, D. A prospective economic analyusis of cost-effectiveness of cetuximab for metastatic colorectal cancer patioents form the NCIC CTG and AGITG Co.17 trial. *Journal of Clinical Oncology* 2008; **26** 6528-

35. Cetuximab in the first line treatment of metastatic colorectal cancer (mCRC). NICE (ERG) Clarification Questions Merck Serono Response. Available at *http://www.nice.org.uk/nicemedia/pdf/ResponseMerckSeronotorequestforclarification.pdf* Accessed 25 September 2009;

36. Bidard, F., Bengrine, L., Figer, A., Cervantes, A., Andre, T., Lledo, G., Mabro, M., Tournigand, C., Louvet, C., de Gramont, Aimery, and GERCOR (French Oncology Research Group) Efficacy of FOLFIRI3 (irinotecan 100mg/m2 D!,D3 combined with LV5FU) or other irinotecan-based regiments in ocaliplatin-pretreated metastatic colorectal cancer ion the OPTIMOX1 study. *Journal of Clinical Oncology* 2008; **26** 4059-

37. Maughan, T. S., James, R. D., Kerr, D. J., Ledermann, J. A., Seymour, M. T., Topham, C., McArdle, C., Cain, D., Stephens, R. J., and Medical Research Council Colorectal Group Comparison of intermittant and continuous palliative chemotherapy for advanced colorectal cancer: a multicentre randomised trial. *Lancet* 2003; **361** 457-464.

38. US National Institutes of Health Combination chemotherapy and bevacizumab with or without erlotinib in treating patients with metastatic colorectal cancer that cannot be removed by surgery. *Available at http://clinicaltrials.gov/ct/show/NCT00265824?order=2.* Accessed Sep 25, 2009.

39. The Oncology Clinical Trials Office (OCTO) QUASAR 2 - A multicentre international study of capecitabine ± bevacizumab as adjuvant treatment of colorectal cancer. *Available at http://www.octo-oxford.org.uk/alltrials/trials/q2.html* Accessed Sep 25, 2009

APPENDICES

Appendix 1. Treatment subgroup comparisons for overall survival and progression free survival (intention to treat analysis)

Interventions (Regimens)[a]	Median follow up (months)	Numbers followed in each group (n)	Number of patients with event (n)	Median progression free survival, (months)	Hazard Ratio (97.5% CI; p-value)
Overall survival					
Initial 2 arm design					
XELOX	-	317	250 (78.9%)	18.8	Not applicable
FOLFOX-4 (control)	-	317	262 (82.6%)	17.7	Not applicable
2x2 factorial design					
B-XELOX	-	350	211 (60.3%)	21.4	Not applicable
P-XELOX	-	350	231 (66.0%)	19.2	Not applicable
B-FOLFOX-4	-	349	209 (59.9%)	21.2	Not applicable
P-FOLFOX-4 (control)	-	351	224 (63.8%)	20.4	Not applicable
Additional analysis requested by the ERG (Post-hoc treatment subgroup comparisons)					
Initial 2 arm design					
XELOX vs. FOLFOX-4					0.90 (0.74, 1.10; p= Not reported)
2x2 factorial design					
B-FOLFOX-4 vs. P-FOLFOX-4					0.94 (0.75, 1.16; p= 0.4937)
B-XELOX vs. P-XELOX					0.84 (0.68, 1.04; p= 0.0698)
Other comparisons					
B-FOLFOX-4 vs. P-XELOX					Not reported
B-FOLFOX-4 vs. B-XELOX					0.99 (0.80, 1.23; p= Not reported)
B-XELOX vs. P-FOLFOX-4					Not reported
P-XELOX vs. P-FOLFOX-4					Not reported
FOLFOX-4 vs. P-FOLFOX-4					Not reported
FOLFOX-4 vs. P-XELOX					Not reported
FOLFOX-4 vs. B-FOLFOX-4					Not reported

Interventions (Regimens)[a]	Median follow up (months)	Numbers followed in each group (n)	Number of patients with event (n)	Median progression free survival, (months)	Hazard Ratio (97.5% CI; p-value)
FOLFOX-4 vs. B-XELOX					Not reported
XELOX vs. P-FOLFOX-4					Not reported
XELOX vs. P-XELOX					Not reported
XELOX vs. B-FOLFOX-4					Not reported
XELOX vs. B-XELOX					Not reported
Progression free survival					
Initial 2 arm design					
XELOX	-	317	290 (91.5%)	7.1	Not applicable
FOLFOX-4 (control)	-	317	299 (94.3%)	7.7	Not applicable
2x2 factorial design					
B-XELOX	-	350	295 (84.3%)	9.3	Not applicable
P-XELOX	-	350	301 (86.0%)	7.4	Not applicable
B-FOLFOX-4	-	349	299 (85.7%)	9.4	Not applicable
P-FOLFOX-4 (control)	-	351	321 (91.5%)	8.6	Not applicable
Additional analysis requested by the ERG (Post-hoc treatment subgroup comparisons)					
Initial 2 arm design					
XELOX vs. FOLFOX-4					0.95 (0.79, 1.15; p= Not reported)
2x2 factorial design					
B-FOLFOX-4 vs. P-FOLFOX-4					0.89 (0.74, 1.06; p= 0.1312) [b]
B-XELOX vs. P-XELOX					0.80 (0.66,0.96; p= 0.0059) [c]
Other comparisons					
B-FOLFOX-4 vs. P-XELOX					0.81 (0.68, 0.98; p= 0.0108)
B-FOLFOX-4 vs. B-XELOX					0.99 (0.82, 1.19; p= Not reported)
B-XELOX vs. P-FOLFOX-4					0.87 (0.73, 1.05; p= 0.0965)
P-XELOX vs. P-FOLFOX-4					Not reported
FOLFOX-4 vs. P-FOLFOX-4					Not reported
FOLFOX-4 vs. P-XELOX					Not reported
FOLFOX-4 vs. B-FOLFOX-4					Not reported

Interventions (Regimens)[a]	Median follow up (months)	Numbers followed in each group (n)	Number of patients with event (n)	Median progression free survival, (months)	Hazard Ratio (97.5% CI; p-value)
FOLFOX-4 vs. B-XELOX					Not reported
XELOX vs. P-FOLFOX-4					Not reported
XELOX vs. P-XELOX					Not reported
XELOX vs. B-FOLFOX-4					Not reported
XELOX vs. B-XELOX					Not reported

NS, not significant

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4, B- alone, Bevacizumab only

[b] value reported as 0.89 (0.73, 1.08; p= 0.1871) in the original peer reviewed paper[5]

[c] value reported as 0.77 (0.63, 0.94; p= 0.0026) in the original peer reviewed paper[5]

Appendix 2. Hazard ratios after step-wise exclusion of subgroups of patients with previous adjuvant chemotherapy in the NO16966 trial

Population	Number of patients excluded from analysis (n)	Number of patients included in analysis (n)	HR (97.5%CI)	p-value
Overall survival				
All patients included (ITT)	0	1400	0.89 (0.76, 1.03)	0.0769
Exclusion of patients with adjuvant chemotherapy from all four treatment arms	85+91+88+76	1060 (1400-340)	0.83 (0.70, 0.99)	0.0183
Exclusion of patients with adjuvant chemotherapy from FOLFOX arms only	85+88	1227 (1400-173)	0.85 (0.72, 1.00)	0.0242
Exclusion of patients with adjuvant chemotherapy from P-FOLFOX arm only	85	1315 (1400-85)	0.84 (0.72;0.98)	0.0116
Progression free survival				
All patients included (ITT)	0	1400	0.83 (0.72, 0.95)	0.0023
Exclusion of patients with adjuvant chemotherapy from all four treatment arms	85+91+88+76	1060 (1400-340)	0.74 (0.64, 0.87)	<0.0001
Exclusion of patients with adjuvant chemotherapy from FOLFOX arms only	85+88	1227 (1400-173)	0.75 (0.65, 0.87)	<0.0001
Exclusion of patients with adjuvant chemotherapy from P-FOLFOX arm only	85	1315 (1400-85)	0.77 (0.67;0.89)	<0.0001

ITT, intention to treat; HR, hazard ratio; P-FOLFOX, placebo plus oxaliplatin plus 5-fluorouracil and folinic acid

Appendix 3. Reasons for stopping treatment (by study arm) during the primary treatment phase of the NO16966 trial (data derived from manufacturer's supplementary evidence)

Reasons for stopping treatment	Individual treatment groups [a] of the initial 2 arm study (n)		Individual treatment groups [a] of the 2x2 factorial study (n)			
	FOLFOX	XELOX	P-FOLFOX	P-XELOX	B-FOLFOX	B-XELOX
Subjects randomised (intention to treat population)	317	317	351	350	349	350
Safety	99 (31.2%)	113 (35.6%)	77 (21.9%)	74 (21.1%)	109 (31.2%)	117 (33.4%)
Abnormality of lab test	0	0	0	0	0	0
Adverse event [b]	91 (28.7%)	99 (31.2%)	72 (20.5%)	72 (20.6%)	101 (28.9%)	109 (31.1%)
Death	8 (2.5%)	14 (4.4%)	5 (1.4%)	2 (0.6%)	8 (2.3%)	8 (2.3%)
Non-safety	203 (64.0%)	182 (57.4%)	237 (67.5%)	235 (67.1%)	188 (53.9%)	179 (51.1%)
Insufficient therapeutic response	127 (40.1%)	131 (41.3%)	155 (44.2%)	174 (49.7%)	102 (29.2%)	101 (28.9%)
Early improvement	0	0	0	0	0	0
Violation of selection criteria at entry	2 (0.6%)	0	3 (0.9%)	4 (1.1%)	5 (1.4%)	1 (0.3%)
Other protocol violation	1 (0.3%)	0	1 (0.3%)	1 (0.3%)	1 (0.3%)	0
Refused treatment [c]	30 (9.5%)	17 (5.4%)	33 (9.4%)	25 (7.1%)	36 (10.3%)	29 (8.3%)
Failure to return	5 (1.6%)	5 (1.6%)	0	0	2 (0.6%)	0
Other	38 (12.0%)	29 (9.1%)	45 (12.8%)	31 (8.9%)	42 (12.0%)	48 (13.7%
Total	302 (95.3%)	295 (93.1%)	314 (89.5%)	309 (88.3%)	297 (85%)	296 (84.6%)

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4

[b] Includes intercurrent illness

[c] Including "did not co-operate"and "withdrew consent"

Appendix 4. Number of patients discontinuing treatment (by study arm) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

	Individual treatment groups [a] of the initial 2 arm study (n)		Individual treatment groups [a] of the 2x2 factorial study (n)			
	FOLFOX-4	XELOX	P-FOLFOX-4	P-XELOX	B-FOLFOX-4	B-XELOX
Number of patients (safety population)	313	316	335	339	342	353
Discontinued treatment due to adverse event						
All grade	91 (29.1%)	99 (31.3%)	68 (20.3%)	72 (21.2%)	105 (30.7%)	109 (30.9%)
Grade 1 and 2 only	NR	NR	NR	NR	NR	NR
Grade 3 and 4 only	NR	NR	NR	NR	NR	NR
Bevacizumab targeted	NR	NR	NR	NR	NR	NR
Treatment-related deaths	7 (2.2%)	10 (3.1%)	7 (2.1%)	6 (1.8%)	7 (2.0%)	8 (2.3%)

NR, not reported

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4

Appendix 5. Grade 3 and 4 adverse events (by study arm) in the NO16966 trial (data derived from manufacturer's supplementary evidence)

Adverse events (grades 3 and 4 unless otherwise stated)	Individual treatment groups[a] of the initial 2 arm study (n)		Individual treatment groups [a] of the 2x2 factorial study (n)			
	FOLFOX-4	XELOX	P-FOLFOX-4	P-XELOX	B-FOLFOX-4	B-XELOX
Number of patients (safety population)	313	316	335	339	342	353
All grade 3 and 4	240 (76.7%)	231 (73.1%)	266 (79.4%)	237 (69.9%)	289 (84.5%)	266 (75.4%)
Grade 4 only	79 (25.2%)	45 (14.2%)	82 (24.5%)	36 (10.6%)	104 (30.4%)	60 (17.0%)
Any related serious adverse event	75 (24.0%)	64 (20.2%)	70 (20.9%)	79 (23.3%)	87 (25.4%)	95 (26.9%)
Gastrointestinal disorders						
All grade 3 and 4	95 (30.4%)	102 (32.3%)	72 (21.5%)	114 (33.6%)	87 (25.4%)	132 (37.4%)
Grade 4 only	6 (1.9%)	12 (3.8%)	4 (1.2%)	12 (3.5%)	7 (2.0%)	11 (3.1%)
Stomatitis	7 (2.2%)	2 (0.6%)	6 (1.8%)	6 (1.8%)	12 (3.5%)	7 (2.0%)
Diarrhoea	40 (11.9%)	63 (19.9%)	34 (10.1%)	70 (20.6%)	44 (12.9%)	77 (21.8%)
Nausea/vomiting	33 (10.5%)	24 (7.6%)	14 (4.2%)	28 (8.3%)	25 (7.3%)	38 (10.8%)
Blood and lymphatic disorders						
All grade 3 and 4	152 (48.6%)	56 (18.0%)	166 (69.3%)	48 (14.2%)	159 (46.5%)	125 (35.4%)
Grade 4 only	58 (17.3%)	6 (1.9%)	66 (19.7%)	8 (2.4%)	73 (21.3%)	44 (12.5%)
Neutropenia/ granulocytopenia	130 (41.5%)	20 (6.3%)	152 (45.4%)	26 (7.7%)	138 (40.4%)	25 (7.1%)
Febrile neutropenia	15 (4.8%)	5 (1.6%)	16 (4.8%)	1 (0.3%)	15 (4.4%)	4 (1.1%)
Gastrointestinal perforation	1 (0.3%)	1 (0.3%)	-	2 (0.6%)	1 (0.3%)	3 (0.8%)
Bleeding problems	5 (16.0%)	7 (2.2%)	2 (0.6%)	6 (1.8%)	7 (2.0%)	6 (1.7%)
Neurological and other toxicity						
Hand/foot syndrome (grade 3)	4 (1.3%)	21 (6.6%)	4 (1.2%)	19 (5.6%)	6 (1.8%)	42 (11.9%)
Neurotoxicity	40 (12.8%)	51 (16.1%)	67 (20.0%)	63 (18.6%)	61 (17.8%)	64 (18.1%)
Venous thromboembolic events	17 (5.4%)	16 (5.0%)	24 (7.2%)	9 (2.7%)	32 (9.4%)	22 (6.2%)
Arterial thromboembolic events	4 (1.3%)	1 (0.3%)	4 (1.2%)	3 (0.9%)	5 (1.5%)	7 (2.0%)
Hypertension	1 (0.3%)	1 (0.3%)	4 (1.2%)	4 (1.2%)	12 (3.5%)	16 (4.5%)
Proteinuria	-	12 (4.7%)	-	-	3 (0.9%)	21 (5.9%)
Wound healing complications	2 (0.6%)	-	2 (0.6%)	-	-	3 (0.8%)
Fistula/intrabdominal abscess	5 (1.6%)	3 (9.5%)	-	1 (0.3%)	4 (1.2%)	2 (0.6%)
Cardiac disorders	8 (2.5%)	4 (1.3%)	1 (0.3%)	2 (0.6%)	23 (6.7%)	14 (4.0%)
Infections/ infestations	35 (11.2%)	26 (8.2%)	31 (9.3%)	19 (5.6%)	30 (8.8%)	21 (5.9%)
Laborotory abnormalities						
Low neutrophils	129 (41.2%)	30 (9.5%)	154 (46.0%)	28 (8.3%)	145 (42.4%)	25 (7.1%)
Low haemoglobin	11 (3.5%)	11 (3.5%)	5 (1.5%)	8 (2.4%)	10 (2.9%)	7 (2.0%)
Low platelets	14 (4.4%)	30 (9.5%)	14 (4.2%)	24 (7.1%)	12 (3.5%)	15 (4.2%)

[a] XELOX, oxaliplatin plus capecitabine; FOLFOX-4, oxaliplatin plus 5-fluorouracil and folinic acid; P-XELOX, placebo plus XELOX; B-XELOX, bevacizumab plus XELOX; P-FOLFOX-4; placebo plus FOLFOX-4; B-FOLFOX-4, bevacizumab plus FOLFOX-4