Pemetrexed for the maintenance treatment of locally advanced or metastatic non-small cell lung cancer (NSCLC)

ERG Report

Contains no commercially in confidence data

Produced by:

Liverpool Reviews and Implementation Group
University of Liverpool
Room B05
Whelan Building
The Quadrangle
Brownlow Hill
Liverpool
L69 3GB
Tel: +44 (0) 151 794 5682/5541/5067
Fax: +44 (0) 151 794 5585
Email: LRiG@liv.ac.uk

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/207/01

Date completed: October 26th 2009

Authors:

Janette Greenhalgh, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool

Claire McLeod, Research Fellow (Health Economics), Liverpool Reviews and Implementation Group, University of Liverpool

Adrian Bagust, Professorial Research Fellow (Health Economics), Liverpool Reviews and Implementation Group, University of Liverpool

Angela Boland, Research Fellow (Health Economics), Liverpool Reviews and Implementation Group, University of Liverpool

Nigel Fleeman, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool

Yenal Dundar, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool

James Oyee, Medical Statistician, Liverpool Reviews and Implementation Group, University of Liverpool

Rumona Dickson, Director, Liverpool Reviews and Implementation Group, University of Liverpool

Helen Davis, Assistant Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

John Green, Consultant Oncologist, Clatterbridge Centre for Oncology NHS Foundation Trust, Liverpool

Emer McKenna, Specialist Registrar in Palliative Care, Willowbrook Hospice, Prescot

Mike Pearson, Professor of Clinical Evaluation and Consultant Respiratory Physician, Aintree University Hospitals NHS Foundation Trust, Liverpool

Declared competing interests of authors: None

Correspondence to: Ms. Rumona Dickson, Director (LRiG), Liverpool Reviews and Implementation Group, University of Liverpool, Room BO5, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

This report should be referenced as follows:
Greenhalgh J, McLeod C, Bagust A, Boland A, Fleeman N, Dundar Y, Oyee J, Dickson R, Davis H, Green J, McKenna E, Pearson M. Pemetrexed for maintenance treatment of locally advanced or metastatic non-small cell lung cancer (NSCLC): A Single Technology Appraisal. LRiG, The University of Liverpool, 2009.

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.

Acknowledgements:
The authors are pleased to acknowledge the help of Catrin Tudor-Smith, Antonieta Medina Lara and Rubin Mujica-Moto who provided statistical and economic advice and commented on drafts of the ERG report.

Contributions of authors:
The review of the clinical evidence was undertaken primarily by Janette Greenhalgh with assistance from Nigel Fleeman (quality assessment) and Yenal Dundar (literature searching). James Oyee provided statistical advice. Helen Davis, John Green, Emer McKenna and Mike Pearson contributed to the critical appraisal of the manufacturer's submission as presented in the clinical sections of the ERG report. Adrian Bagust carried out the critical appraisal of the manufacturer's economic model with assistance from Claire McLeod and Angela Boland. Claire McLeod summarised the manufacturer's review of economic literature and described the economic model. All authors read and commented on draft versions of the ERG report.

Contents

Abbreviations:

AE(s)	Adverse event(s)
ASCO	American Society of Clinical Oncology
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CSR	Clinical study report
CTX	Chemotherapy
ECOG	Eastern Cooperative Oncology Group
EMEA	European Medicines Agency
EQ-5D	EuroQol 5D (a standardised instrument used as a measure of health outcome)
ERG	Evidence Review Group
HEED	Health Economic Evaluation Database
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient data
ITT	Intention to treat
iv	Intravenous
LUCADA	Lung Cancer Data
LCSS	Lung Cancer Symptom Scale
LYG	Life year gained
MS	Manufacturer submission
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Clinical Excellence
NSCLC	Non-small cell lung cancer
NSCLC-NOS	Non-small cell lung cancer not otherwise specified
OS	Overall survival
PFS	Progression free survival
PP	Per protocol
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
SA	Sensitivity analysis
SPC	Summary of Product Characteristics
STA	Single Technology Appraisal
WTP	Willingness to pay

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE from Eli Lilly in support of the use of pemetrexed (Alimta®) for the maintenance treatment of non-small cell lung cancer (NSCLC) within its current licence. The manufacturer submission (MS) describes the use of pemetrexed as maintenance therapy for patients whose disease has not progressed following the completion of four cycles of first-line (induction) chemotherapy (CTX).

In July 2009, the Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency (EMEA), approved an extension to the licence for the use of pemetrexed (Alimta®) "as monotherapy for the maintenance treatment of patients with NSCLC, other than predominantly squamous cell histology. First-line treatment should be a platinum doublet with gemcitabine, paclitaxel or docetaxel."[1]

1.2 Summary of submitted clinical-effectiveness evidence

The evidence described in the MS is derived from a double-blind, placebo-controlled randomised controlled trial (RCT), the JMEN trial.[2] The trial compared the use of pemetrexed + best supportive care (BSC) as maintenance therapy with placebo + BSC in patients with NSCLC (n=663) who had received four cycles of platinum-based CTX and whose disease had not progressed. The MS focussed on the clinical outcomes of the subgroup of patients with non-squamous histology (n=481) which is the population for which pemetrexed is licensed in this indication; the MS also focussed on a subgroup of the licensed population, patients with adenocarcinoma.

In the licensed non-squamous population, the trial demonstrated greater median progression free survival (PFS) for patients treated with pemetrexed compared to patients in the placebo arm (4.5 vs. 2.6 months; HR 0.44; 95% CI 0.36-0.55, p<0.00001). Median overall survival (OS) was also greater for the pemetrexed-treated patients (15.5 vs. 10.3 months; HR 0.70; 95% CI 0.56-0.88, p=0.002). In addition, tumour response and disease control rates were statistically significantly greater for patients who received pemetrexed. Patient survival rates at one year and two years were higher in the pemetrexed arm. The health-related quality of life (HRQoL) data presented were limited due to high levels of censoring/missing data. Safety

data demonstrated that patients treated with pemetrexed had statistically significantly higher rates of grade 3 or 4 neutropenia, experienced higher rates of transfusions and hospitalisation due to drug toxicity.

1.3 Summary of submitted cost-effectiveness evidence

The manufacturer did not identify any published cost-effectiveness analyses of pemetrexed for the maintenance treatment of patients with NSCLC, and therefore developed a *de novo* economic model to support their economic case. The model compares pemetrexed + BSC with 'watch and wait' + BSC. The clinical data used in the economic model were primarily generated from the JMEN trial.[2] Although the model was trial-based, there was also a modelling component to allow the extrapolation of health effects beyond the 29 month trial period up to six years. The manufacturer's economic evaluation adopts a lifetime horizon (taken as six years) for the consideration of costs and benefits and the perspective is that of the UK NHS and Personal Social Services (PSS).

The ICERs estimated by the manufacturer's model are £33,732 per QALY for the licensed non-squamous population, and £39,364 per QALY for the adenocarcinoma subgroup. Both of these ICERs are above the standard NICE willingness to pay range (£20,000-£30,000 per QALY).

The manufacturer has presented a case for pemetrexed to be considered as an end of life treatment.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

The manufacturer cites evidence from a well-designed trial (JMEN[2]) of the clinical benefit of pemetrexed + BSC as maintenance treatment compared with placebo + BSC. The trial recruited a substantial number of patients in a difficult disease area. It is noteworthy that patients and assessors in the JMEN[2] trial were blinded to treatment group allocation and that investigators' outcome assessments were independently verified.

1.4.2 Weaknesses

Clinical

The ERG notes that there is only one relevant RCT (JMEN[2]) which compares pemetrexed + BSC as maintenance treatment with placebo + BSC. Despite designing the trial to include a comprehensive analysis of HRQoL, very limited data were collected and reported in the MS.

This means it is very difficult to determine how patients' HRQoL will be affected by pemetrexed in a maintenance setting.

The primary endpoint of the key trial was changed by the manufacturer from OS to PFS during the course of the trial. No information was provided that fully justified the change of clinical endpoint and it is not clear at what time point the decision was made. The statistical analysis plan described by the manufacturer also included a test for treatment by histology interaction and corresponding subgroup analyses. The results for the subgroup of patients with non-squamous histology provide the clinical evidence in the MS. However, the trial randomisation process did not include stratification by histology status and the trial was not powered to perform the subgroup analysis, thus reliance on the results should be treated with due caution. Moreover, the restriction of the licensed population to only the non-squamous subgroup effectively reduces the statistical power of the trial, with consequences of increased uncertainty in the cost-effectiveness analysis.

Economic

The projection of survival from the end of the trial period, the costing of CTX treatment and the utility values used in the manufacturer's model are not ideal and underestimate the size of the ICER.

The manufacturer implemented a capping rule in their economic model to limit the maximum number of cycles of maintenance treatment that patients could receive. However, the cycle capping rule only affects costs; it does not take account of any reduction in outcomes caused by capping the maximum number of cycles at 17 rather than allowing the JMEN trial[2] maximum of 55. Again, this capping rule underestimates the size of the ICER.

Making all of the necessary ERG corrections/adjustments to the manufacturer's model, the ERG's base case ICER for the non-squamous population is estimated at £51,192 per QALY.

1.5 Areas of uncertainty

1.5.1 Is the evidence generalisable to UK practice?

The generalisability of the JMEN trial[2] to UK clinical practice is uncertain for a number of reasons:

- None of the patients in the trial were recruited from the UK. A sizeable proportion (35%) of patients were from Asian countries; these patients are documented in the literature as having a better prognosis for NSCLC than other ethnic groups and the Asian patients in the trial appear to have improved survival times compared with patients of other ethnicities.[3]

- Patients in the trial were able to receive unlimited cycles of maintenance therapy. This is unlikely to be the case in clinical practice in England and Wales and it is unclear how this difference would impact on survival in a clinical setting.

- The trial excluded patients who had received pemetrexed or vinorelbine as a first-line treatment; hence there is no information on how patients treated with first-line vinorelbine or pemetrexed will respond to pemetrexed administered as maintenance therapy. These patients will therefore not be eligible for pemetrexed maintenance therapy.

- Paclitaxel appeared to be used as a first-line treatment for a greater proportion of patients in the trial than might be the case in clinical practice in England and Wales. The impact of this when generalising the results to clinical practice in England and Wales is unknown.

- A number of patients in the trial received second-line therapies that are not available to patients in clinical practice in England and Wales, which may have affected the OS observed in the trial.

A further area of uncertainty is that confirmed histological diagnosis of non-squamous NSCLC is required before patients can be offered maintenance treatment with pemetrexed. Whilst histological testing is routinely carried out in many centres in England and Wales, this will not be available to all patients. Therefore, it is unclear if pemetrexed for maintenance therapy will be available in all centres in the UK, which may give rise to equity concerns.

1.5.2 'End of Life' criteria

Analysis of the JMEN trial[2] individual patient data (IPD) and revised projection modelling confirms that the mean life extension from use of pemetrexed as maintenance therapy is likely to exceeed 3 months.

However, the number of patients who would be eligible to receive pemetrexed is uncertain. The manufacturer's estimates (used to present their end of life case) are based on amalgamation of information from different sources with differing definitions. The methods of calculation are not well reported and a number of assumptions have been made which may not be valid.

1.6 Summary

Several factors serve to limit the generalisability of the trial to UK clinical practice, and the ERG cannot be confident that the clinical results presented in the MS give a true reflection of the benefits that could be expected with pemetrexed for the maintenance treatment of patients with non-squamous NSCLC in UK clinical practice. Furthermore, in the economic analysis there were a number of problems identified with the model (in addition to the JMEN trial[2] data) which indicate that the ICER (re-estimated as £51,192 per QALY gained) could well exceed NICE's willingness to pay thresholds.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

In the context section of the MS (section 4 of MS), the manufacturer describes the key issues relating to lung cancer (see Box 2-1). The ERG considers the description to be an accurate account.

Box 2-1 Summary of the manufacturer's description of lung cancer

Lung cancer is the second most common cancer diagnosed in the UK, with over 33,000 new cases diagnosed in England and Wales in 2006 and the leading cause of cancer death.[4] Lung cancer is the second most common cancer in men after prostate cancer, and the third most common cancer in women after breast and bowel cancer.

Non-small cell lung cancer (NSCLC) accounts for approximately 80% of all lung cancers diagnosed. The main sub-types of NSCLC are squamous cell carcinoma (33%), adenocarcinoma (25%), large cell carcinoma (4%), and 36% being NSCLC 'not-otherwise specified' (NOS).[5] While cigarette smoking has been linked to all four types of lung cancer, the incidence of adenocarcinoma has been steadily increasing worldwide, and modifications to cigarette design are thought to be responsible for this shift in pathologic diagnosis pattern.[6]

Survival in patients with lung cancer is poor. It was responsible for approximately 29,600 deaths in England and Wales in 2007.[4] For patients with stage IIIB, only 7-9% may live for 5 years and for patients with stage IV (metastatic) cancer, only about 2-13% survive for 5 years[4]

One reason for this poor prognosis is the late identification of the disease. Lung cancer is asymptomatic in the early stages and advanced disease is not amenable to curative treatment. Another reason, which explains the UK's relatively poor performance in comparison with other developed countries, is low active anti-cancer treatment rates. The National Lung Cancer Audit states that only 23.2% NSCLC patients in England and Wales received first-line chemotherapy in 2006.[7]

2.2 Critique of manufacturer's overview of current service provision

The MS describes the role of maintenance therapy (see Box 2-2) and establishes it as a new treatment option between first- and second-line therapy. The ERG accepts this definition of maintenance therapy as appropriate.

Box 2-2 Summary of the manufacturer's description of maintenance therapy

In accordance with the licence for pemetrexed, in this appraisal maintenance treatment is defined as the administration of additional chemotherapy immediately after the completion of first-line (induction) chemotherapy in patients with complete / partial response or stable disease (as defined by RECIST criteria) after four cycles of induction chemotherapy. Patients who have disease progression after induction treatment are not eligible for maintenance treatment.

The goal of maintenance treatment is to maintain the clinical benefit achieved with first-line chemotherapy. Maintenance treatment is continued until disease progression.

Maintenance treatment is a new treatment paradigm and is proposed as an alternative for the **'watch and wait'** phase of the current treatment pathway, for patients with complete or partial response / stable disease after four cycles of first-line treatment.

As part of its rationale for maintenance treatment, the manufacturer asserts that such treatment is 'routinely given in current clinical practice for other cancers like breast cancer, lymphoma and prostate cancer' (MS, pg18). However the ERG notes that the comparison between NSCLC and other cancers may not be valid since patients with NSCLC generally present later, are less likely to benefit from surgical interventions than other solid tumours, and usually have a poorer prognosis.

The MS provides a diagram outlining the current treatment pathway and the proposed place for pemetrexed maintenance therapy which the ERG regards as appropriate. However, pemetrexed is only licensed for non-squamous NSCLC, which means that patients will require histological diagnosis before receiving pemetrexed. The manufacturer provides LUCADA data[7] to show that 68% of patients (range 20%-85%) in England and Wales had a histological diagnosis of their lung cancer. The ERG's communications with clinical experts confirmed that although histological diagnosis is becoming standard practice, not all centres will be able to offer it at the present time; thus, there may be regional variation and equity concerns.

The ERG further notes that current NICE guidance[8] for first-line CTX treatment also includes vinorelbine in addition to gemcitabine, docetaxel and paclitaxel; vinorelbine is not cited in the manufacturer's diagram. Pemetrexed has recently been approved for first-line use and does not appear in the manufacturer's diagram of the treatment pathway (MS, Figure 1).[9]

In addition, it is not specified in the diagram that only those patients with NSCLC who are considered to be of good performance status (PS 0 or 1) should be offered CTX in clinical practice in England and Wales.[10]

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

The final scope issued by NICE and the manufacturer's definition of the decision problem are described in the MS (pg 8-9) and the summary table is reproduced here.

Table 3-1 Final scope issued by NICE and the manufacturer's definition of the decision problem

	Final scope issued by NICE	Decision problem addressed in the MS
Population	People with advanced or metastatic (stage IIIB and IV) NSCLC, other than those with predominantly squamous histology, whose disease has not progressed following treatment with platinum-based, first-line chemotherapy	Patients with locally advanced or metastatic NSCLC of other than predominantly squamous (non-squamous) histology whose disease has not progressed [i.e., have complete response (CR), partial response (PR) or stable disease (SD)] following four cycles of induction treatment with a platinum doublet (one of the following: gemcitabine, docetaxel, paclitaxel in combination with cisplatin or carboplatin).The base case population for this submission is the licensed population: patients with non-squamous NSCLC (adenocarcinoma, large cell carcinoma or NSCLC 'not otherwise specified').
Intervention	Pemetrexed (maintenance treatment)	Pemetrexed (500mg/m^2 iv infusion) administered on day 1 of a 21-day cycle, until disease progression.
Comparator(s)	Best supportive care, which may include palliative radiotherapy and corticosteroids (without maintenance therapy)	Placebo (watch and wait). Both treatment arms received BSC
Outcomes	• health-related quality of life • overall survival • progression free survival • response rates • adverse effects of treatment	• health related quality of life • overall survival • progression free survival • response rates • adverse effects of treatment
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 perspective.	Cost-effectiveness analysis results expressed as incremental cost per QALY gained. A cost per Life Year (cost per LY) gained analysis will also be conducted as this type of analysis is relevant in disease areas where extended survival is a key outcome of treatment. The time horizon is six years, (a lifetime model). Costs will be considered from an NHS and Personal Social Services perspective.
Subgroups to be considered		Results for the non-squamous patients (licensed population) and for the subgroup of patients with adenocarcinoma are presented in this submission.
Special considerations		None

3.1 Population

The stated population in the final scope issued by NICE and discussed in the manufacturer's definition of the problem is patients with advanced or metastatic (stage IIIB and IV) NSCLC, other than those with predominantly squamous histology, whose disease has not progressed following treatment with platinum-based, first-line CTX. In the MS, the manufacturer's clinical evidence is only applicable to those patients who have received a first-line platinum doublet containing gemcitabine, paclitaxel or docetaxel (as reflected by the licensed indication); this means that no inference about the clinical effectiveness of pemetrexed maintenance in patients who received first-line pemetrexed or vinorelbine can be made.

3.2 Intervention

Pemetrexed (Alimta®) is given as an intravenous (iv) infusion at a dose of 500mg/m^2 on day 1 of a 21 day cycle until disease progression. The JMEN[2] clinical trial, which provides the evidence-base for the MS and the key licensing information, placed no limits on the maximum number of CTX cycles administered to patients. The manufacturer notes that in clinical practice a maximum of 15-20 cycles is likely to be administered to patients; however, the ERG considers that since maintenance treatment for NSCLC is new to the NHS the number of treatment cycles is unknown. Patients also receive BSC (as needed) alongside treatment with pemetrexed.

3.3 Comparators

The stated comparator in the final scope is 'watch and wait' plus BSC as needed. In the definition of the decision problem, the manufacturer's stated comparator is placebo (watch and wait) plus BSC. The placebo treatment consisted of a saline solution infused over ten minutes. Therefore, the ERG considers that the manufacturer's comparator includes an extra element to the 'watch and wait' policy used in clinical practice. Whilst the placebo treatment adds to the robustness of the trial design, in clinical practice placebo treatment would not be offered.

3.4 Outcomes

The MS identifies OS, PFS, HRQoL, response rates and adverse events (AEs) as key outcomes, which match those within the scope issued by NICE and are standard for research in this field.

3.5 Time frame

In the JMEN RCT,[2] which is the key source of clinical data, patients were followed up until death or study closure. At the time of data lock the maximum duration of censored survival was 38 months. The economic model uses a six year time frame, which is taken to be equivalent to a life-time horizon.

3.6 Other relevant factors

No specific subgroup analyses are defined in the NICE scope; however, the manufacturer has identified patients with adenocarcinoma histology as an important subgroup of patients from within the non-squamous population.

4 CLINICAL EFFECTIVENESS

Table 4-1 provides an outline of the key background/clinical information and its location within the MS. Its purpose is to signpost the reader to the main areas of background/clinical information within the MS.

Table 4-1 Key non-economic information in the MS

Key information	Pages in the MS	Key tables/figures in the MS
Description of technology	3-7	
Statement of decision problem	8-9	
Context/background	15-21	Figure 1, page 16
Equity and equality	22	
Literature search:		
Search strategies	Appendix 2	
Study selection	23-24	
Clinical effectiveness evidence:		
Trial information	25-43	Table 2, page 30
Results: main and subgroups	43-45, 47	Table 8, page 47
Results: HRQoL analysis	45-46	Table7, page 46
Results: end of life	48-49	Table 10, page 49
Results: safety	50-53	Table 12-13, page 52

4.1 Critique of manufacturer's approach

4.1.1 Description of manufacturer's search strategy and comment on the appropriateness of the chosen search strategy.

The stated aim of the literature search described in the MS was to identify studies of pemetrexed maintenance in patients with advanced NSCLC. A 'broad-based' search strategy was implemented given that maintenance treatment is a relatively new concept and that earlier literature searches had revealed i) that studies often report combined outcomes following first-line and maintenance treatments, and ii) the term 'maintenance' treatment may not be consistently interpreted by different investigators (MS, pg27).

The search was comprehensive and included appropriate databases: BIOSIS Previews 1989 to 2009 Week 24, Current Contents/All Editions1993 Week 27 to 2009 Week 22, EMBASE 1980 to 2009 Week 20, Ovid MEDLINE(R) 1950 to May Week 3 2009, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations May 21, 2009, Cochrane DSR, ACP Journal Club,

DARE, CCTR, CMR, HTA, and NHSEED. Searches were also conducted of the American Society of Clinical Oncology conference website (ASCO) and of databases held by Eli Lilly.

The search strategy described in Appendix 2 of the MS used a filter to identify RCTs and combined drug names with disease. The ERG notes that part of the filter was not incorporated into the final search string; however, this is considered a minor error.

The ERG undertook its own searches and is confident that all relevant trials were identified by the manufacturer.

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

Table 4-2 shows the inclusion and exclusion criteria presented in the MS. The ERG considers the inclusion/exclusion criteria to be appropriate given the manufacturer's stated objectives.

Table 4-2 Inclusion and exclusion criteria

Inclusion	Exclusion
Randomised controlled trials	Phase I/II
Phase III	First-line NSCLC only
Pemetrexed in maintenance treatment of advanced (stage IIIB/IV) NSCLC	Second-line NSCLC only
Head to head comparisons versus pemetrexed	
English language	

The MS lists four relevant articles relating to one RCT; the JMEN trial[2] was the only trial included in the systematic review. Three of the identified articles were abstracts reporting the JMEN trial[2] which were presented at ASCO in 2008; the fourth was the clinical study report for the trial. The MS confirms that at the time of submission, the full text article describing the JMEN trial[2] was not published. The full text article was subsequently published in September 2009.[2]

Table 4-3 JMEN trial characteristics

Trial design and number of participants	Intervention/Comparator	Inclusion criteria (main)	Exclusion criteria (main)	Outcomes
International, multi-centre, Phase III, double-blind, placebo-controlled, randomised controlled trial. 83 centres in 20 countries. Patients were randomised on a 2:1 basis. Of the total n=663 participants, 481 represented the non-squamous (licensed) population	Intervention: Pemetrexed 500mg/m^2 infusion over 10 minutes on day 1 every 21 days plus BSC Comparator: Placebo (normal saline) infusion over 10 minutes on day 1every 21 days plus BSC Patients in both groups received prior and concomitant medication with folic acid, vitamin B_{12}, and dexamethasone as recommended in the pemetrexed SPC Best supportive care defined as treatment without a specific antineoplastic regimen and treatment was administered as considered appropriate by the prescribing physician. Acceptable therapies included, but were not limited to antibiotics, antiemetics, thoracentesis, pleurodesis, blood transfusions, and/or nutritional support. Specific exclusions: anticancer surgery, immunotherapy, radiotherapy, anticancer hormonal therapy, and	• histologic or cytologic diagnosis of NSCLC Stage IIIB (with pleural effusion and/or positive supraclavicular lymph nodes) or Stage IV, prior to induction therapy • received only one of the following induction therapies, based on 21-day cycles and lasting precisely four cycles: gemcitabine plus carboplatin, paclitaxel plus carboplatin, docetaxel plus carboplatin, gemcitabine plus cisplatin, paclitaxel plus cisplatin, or docetaxel plus cisplatin • documented evidence of a tumour response of CR, PR, or SD • ECOG PS of 0 or 1 • at least 18 years of age • adequate organ function • prior radiation therapy was allowed to <25% of bone marrow. Prior radiation to the whole pelvis was not allowed • prior radiotherapy must have been completed at least four weeks	• had received prior systemic anticancer therapy excluding those listed in the inclusion criteria) • had received treatment within the last 30 days with a drug that had not received regulatory approval for any indication at the time of study entry • inability to comply with protocol or study procedures • had a serious concomitant systemic disorder • had a serious cardiac condition • CNS metastases • presence of clinically detectable (by physical exam) third-space fluid collections • concurrent administration of any other antitumour therapy • inability to interrupt aspirin or other NSAIDs for a 5-day period	Primary: PFS Secondary: OS TWS Objective tumour response rate Adverse events Changes in LCSS

	systemic CTX in which the goal would be to either eradicate or slow the progression of the study disease	before study enrolment • signed informed consent document on file • male and female patients with reproductive potential must have been on an approved contraceptive method, if appropriate • estimated life expectancy of at least 12 weeks • patient compliance and geographic proximity that allowed adequate follow-up • patient must have received on-study therapy no earlier than 21 days and no later than 42 days from day 1 of their last cycle of induction therapy	• inability or unwillingness to take folic acid or vitamin B12 supplementation • inability or unwillingness to take corticosteroids • received an induction CTX regimen that was not based on a 21-day cycle • pregnant or breast feeding • a prior malignancy other than NSCLC	

BSC=best supportive care; CNS = central nervous system; CR = complete response; ECOG=Eastern Co-Operative Oncology Group; LCSS= Lung Cancer Symptom Scale; NSAID= non-steroidal anti-inflammatory drugs; NSCLC= non-small cell lung cancer; OS= overall survival; PFS = progression free survival; PR = partial response; PS= performance status; SD = stable disease; SPC = summary of product characteristics; TWS = time to worsening of symptoms

4.1.3 Relevant studies that were not included in the submission

No relevant trials were excluded from the analysis and the ERG is confident that there are no other studies relevant to the review.

4.1.4 Description and critique of manufacturer's approach to validity assessment

Since the submission is based on a single RCT (JMEN[2]), the remainder of this section documents the ERG's assessment of the trial and its applicability to clinical practice in the UK.

The validity assessment carried out by the manufacturer (and reviewed by the ERG in Appendix 1) demonstrated that the JMEN trial[2] was reasonably well-designed. However, the ERG has a number of concerns regarding the conduct of the trial as well as its generalisability to the clinical population of England and Wales.

In the JMEN trial,[2] 663 patients were randomised at 83 centres in 20 countries not including the UK. Participating centres were located in Australia, Austria, Brazil, Bulgaria, China, Croatia, Czech Republic, Germany, Greece, Hungary, India, Italy, Korea, Netherlands, Poland, Rumania, Spain, Taiwan, Turkey and USA. The ERG requested from the manufacturer the results of any analyses carried out to assess trial outcomes by region for the licensed non-squamous population. Table 10-2 (Appendix 2) illustrates the (unadjusted) outcomes for three 'regions'; European Union (EU) (n=230), non-Asian (n=310) and Asian (n=171). These results are based on post-hoc subgroup analyses and should be viewed with caution. However, it can be seen that the relative difference in OS for each region is similar to that of the non-squamous group as a whole. It is notable that patients in the Asian region who represent 35% of the trial population recorded substantially longer absolute OS times than either the EU or non-Asian patients. There is evidence to suggest that this ethnic group has a more favourable prognosis for OS in NSCLC in general and thus may be different to the majority of patients treated in England and Wales.[3, 11] However, key effectiveness results depend on relative differences (rather than absolute values) which do not appear to be affected by ethnicity.

The population in the JMEN trial[2] was restricted to patients with an ECOG PS of 0 or 1 and with few co-morbidities. The ERG notes that healthier and younger patients do participate in clinical trials as inclusion criteria are designed to restrict patient entry in order to limit confounding factors. However, the ERG's communications with clinical experts confirmed that patients with good PS and health status are a relatively small proportion of the total number of NSCLC patients treated in clinical practice in England and Wales.

The manufacturer compares the proportions of patients with each subtype of NSCLC in the JMEN trial[2] to the proportions recorded in LUCADA.[7] The manufacturer suggests that the difference may be explained by better histological diagnosis in the setting of the clinical trial (MS, pg33). The ERG also suggests that these differences might be further explained by the inclusion/exclusion criteria of the trial which limit patient entry to those who are fit and have few co-morbidities.

As noted in the background section, identifying the licensed population requires a more specific histological diagnosis than is currently available across all treatment centres. The MS reports the results of an independent review of a sample of the trial investigators' histology classifications (MS, pg29). The agreement between the two was 89.2%, which the manufacturer claims confirms the high diagnostic accuracy in histological diagnosis. However, the ERG notes that it is unclear from the MS how the disagreements were handled.

In justifying the applicability of the results to the UK patient population, the MS contends that patients received induction regimens similar to that of the average NSCLC patient in the UK (MS, pg42). The ERG notes that the induction therapies in the trial were, in part, similar to those used in clinical practice in England and Wales; however, a substantially larger proportion of patients in the trial received paclitaxel than is likely to be the case in England and Wales. According to the manufacturer's survey of UK oncologists (presented in 2008 in support of a submission for pemetrexed to be used as a first-line treatment for NSCLC) only 1% of patients were treated with paclitaxel.[12] Of the patients in the JMEN trial,[2] 32% received paclitaxel as induction CTX.

In addition, the ERG is aware that patients in clinical practice in England and Wales will also be treated with vinorelbine or pemetrexed as a first-line therapy. None of the patients in the JMEN trial[2] received these treatments (pemetrexed was not licensed for first-line therapy at the start of the JMEN trial[2]) and therefore no inference can be made regarding the efficacy or safety of pemetrexed as a maintenance therapy for patients receiving these first-line treatments.

When the ERG examined the range of second-line therapies used in the JMEN trial,[2] (provided by the manufacturer in response to the ERG clarification request) it was found that 53% of treatments used in the pemetrexed arm of the trial are not used in clinical practice in England and Wales. This was also true for 36% of the second-line treatments given to placebo patients. These treatments may have influenced the OS estimates observed in the trial and may mean the results are not reflective of the survival benefits that might be expected in UK practice.

4.1.5 Description and critique of manufacturers outcome selection

The outcome measures presented in the MS are shown in Table 4-4. These are standard outcomes for a trial of this type and match those specified in the scope. Health-related QoL was assessed using the Lung Cancer Symptom Scale (LCSS). The LCSS is designed as a disease-specific measure of QoL, particularly for use in clinical trials. It evaluates six major symptoms associated with lung malignancies and their effect on overall symptomatic distress, functional activities, and global quality of life.[13]

Table 4-4 Outcome measures included in the JMEN trial

Outcome	Definition & Measure	Timing of assessment
Progression free survival (Primary)	Time from randomisation to the first radiologic confirmation of disease progression, or death from any cause RECIST – based on computed tomography (CT), including spiral CT, scans and magnetic resonance imaging (MRI), or in some cases chest X-rays (when lesion is clearly defined and surrounded by aerated lung)	Baseline (post-induction therapy): after four cycles of treatment and no more than 42 days after last dose of induction therapy On study: repeated every 2 cycles of therapy. Assessment within 7 days prior to day 1 of each cycle Post-study follow-up: for patients without documented disease progression, approximately every 6 weeks until documented objective disease progression. After disease progression, every 90 days until death or study closure
Overall survival	Time from day of randomisation to death from any cause	N/A
Time to worsening of symptoms	Date from randomisation to first date of worsening for each of the six LCSS symptoms and three summary items Worsening defined as a 15mm increase on the 100mm visual analogue scale. For each patient who was not known to have had a worsening (defined in this way), time to worsening of symptoms was censored at the date of the patient's last LCSS assessment. LCSS	Baseline, prior to randomisation, once every cycle until study discontinuation and within 30 days of discontinuation
Objective tumour response rate	Proportion of patients per study arm with a confirmed partial response or complete response RECIST	Baseline (post-induction therapy): after four cycles of treatment and no more than 42 days after last dose of induction therapy On study: Repeated every 2 cycles of therapy. Assessment within 7 days prior to day 1 of each cycle Post-study follow-up: For patients without documented objective disease progression, approximately every 6 weeks until documented objective disease progression Response confirmation: Responding patients must have had confirmatory scans performed within 6 weeks (but not less than 28 days) of the last scan. Responding pts were followed every 6 weeks (but not less than 28 days) until documented disease progression
Adverse events	Rated using the NCI CTC AE scale (Version 3.0; NCI 2003)	On study: repeated every 2 cycles of therapy. Assessment within 7 days prior to day 1 of each cycle

LCSS= Lung Cancer Symptom Scale; NCI CTCAE = National Cancer Institute Common Terminology for Adverse Events; RECIST= Response Evaluation Criteria in Solid Tumours

4.1.6 Description and critique of the statistical approach used

As the systematic review only found one study describing pemetrexed as maintenance therapy, no meta-analysis was undertaken.

The MS documents that the primary endpoint for the JMEN trial[2] was changed from OS to PFS (MS, pg 36) but claims that the design of the JMEN trial[2] allowed for robust evaluation of PFS without requiring any changes in sample size, efficacy assumptions or statistical power of the final OS analysis.

The manufacturer provided the following reasons for changing the primary outcome;

- Patients with lung cancer are now living longer and receiving multiple lines of treatment with the potential to confound interpretation of OS. Thus, PFS may provide a better indicator of the effectiveness of pemetrexed in this setting.

- The double-blind placebo controlled study design of JMEN[2] allowed for robust evaluation of PFS without requiring any changes in sample size, efficacy assumptions or statistical power of the final overall survival analysis.

The ERG considers that the manufacturer's justification for changing the primary endpoint is incomplete and inadequate without sufficient scientific evidence and it is not clear when the decision to change was taken. The primary endpoint is the variable that is capable of providing the most clinically relevant and convincing evidence. However, in some instances, there may be good justification for changing the primary endpoint, for example, if external information suggests that another variable may be better suited to measure patient benefit.

According to the JMEN trial[2] clinical study report (CSR) the decision to change the primary endpoint was taken five months before the last patient completed the study and nine months before the JMEN database was locked and the un-blinding of the study database (CSR, pg650). The point at which the primary outcome was changed appears to be after an interim analysis even though the protocol states that an interim analysis was not planned.

A change in the primary endpoint after an interim analysis should not be acceptable: from an experimental design perspective practicability and efficiency play important roles when electing a primary endpoint at the planning stage of a clinical trial. Nevertheless, justification of a primary endpoint (or a change to a primary endpoint) should be focussed on clinical interpretation: endpoints

are usually not selected based on their ability to differentiate between treatment and control, but rather to describe a relevant clinical benefit in the treatment of the condition under study.[14]

It is stated in the JMEN CSR that the JMEN Statistical Analysis Plan was updated (prior to the JMEN database lock and un-blinding of the study database) to include a pre-specified test for treatment-by histology interaction and corresponding subgroup analyses. The ERG notes that the randomisation in the JMEN trial[2] was not stratified by histology status and the trial was not powered to perform the treatment-by histology interaction.

The trial was conducted in 83 centres across 20 countries and although admirable numbers were recruited, randomisation was applied centrally rather than within centre. Uniformity of general clinical practice within so many centres is not guaranteed. With so many investigators in different countries, general clinical practice will always be an issue and the results of a trial can only be generalisable if it is executed efficiently. The manner in which the protocol is implemented should be clear and similar at all centres if potential clustering is to be minimized.

4.1.7 Summary statement

The systematic review in the MS, which identified only one trial comparing pemetrexed + BSC as maintenance therapy to placebo + BSC was complete and reasonable. The search strategy was appropriate and reasonably reported. All relevant clinical trials were identified and validity of the one (unpublished) included trial (JMEN[2]) was discussed by the manufacturer. The trial was reasonably well-designed, incorporating blinding, placebo and independent monitoring of investigator assessments. The clinical outcomes reported in the single relevant RCT identified cover the relevant outcomes outlined in the final scope issued by NICE (OS, PFS, tumour response, AEs and HRQoL). However, the ERG considers the manufacturer has not properly justified the decision to change the primary endpoint of the JMEN trial[2] from OS to PFS. This decision had the effect of truncating the data available for analysis for OS, which is of critical importance to the economic evaluation. Further, the main evidence furnishing the MS is limited since it is derived from a subgroup of patients who cannot be considered to have been truly randomised. The data relating to HRQoL were limited. The ERG also identified a number of issues that may constrain the generalisability of the trial results to clinical practice in England and Wales; these relate to the lack of UK centres in the trial, the large proportion of Asian patients in the trial and key differences in the types of induction and second-line therapies offered to patients in the trial.

4.2 Summary of submitted clinical-effectiveness evidence

4.2.1 Summary of JMEN trial results

The clinical-effectiveness evidence described in the MS is derived from a phase III, multi-centre, double-blind, placebo-controlled RCT which compared pemetrexed + BSC as maintenance therapy with placebo + BSC. The JMEN trial[2] included 663 patients with either squamous or non-squamous NSCLC. At randomisation, all patients had received four cycles of induction CTX (based on 21-day cycles) and had no documented disease progression. Randomisation took place between 21 and 42 days from day 1 of the last cycle of induction therapy.

The evidence-base for this appraisal is the subgroup of patients with non-squamous histology (the licensed population) which included 481 patients from the overall trial population. As noted earlier, the ERG emphasises that this is a pre-specified subgroup analysis rather than results based on a randomised group. The baseline characteristics of these patients appear to be well balanced between the treatment and placebo arms.

The outcomes for the largest subgroup of patients within the licensed population, patients with adenocarcinoma (non-squamous) were also presented by the manufacturer.

The results for the primary and secondary efficacy outcomes for the non-squamous population of the JMEN trial[2] are outlined in Table 4-5. Patients in the pemetrexed arm had statistically significantly longer median PFS and OS compared to patients in the placebo arm. The tumour response and disease control rates for patients in the pemetrexed and placebo arms are statistically significantly different in favour of pemetrexed. Survival rates at one year are substantially greater in the pemetrexed group compared with patients in the placebo arm. At two years, this difference is minimal.

Table 4-5 Key results of the JMEN trial (non-squamous population)

Endpoint	Pemetrexed (n = 325)	Placebo (n = 156)	HR (95% CI)	p value
Primary				
PFS (months) median	4.5	2.6	0.44 (0.36-0.55)	<0.00001
Secondary				
OS (months) median	15.5	10.3	0.70 (0.56-0.88)	0.002
Tumour response (%) (CR + PR)	7.4	1.9		0.018
Disease control rate (%) (CR+PR+SD)	57.7	32.7		<0.001
Survival rate at 1 year (%)	60	42		
Survival rate at 2 year (%)	28	22		

CI=confidence interval; CR=complete response; HR=hazard ratio; OS=overall survival; PFS=progression free survival; PR=partial response; SD= stable disease

4.2.2 Patients with adenocarcinoma histology

The key clinical effectiveness results for this group of patients are depicted in Table 4-6. Median PFS and OS are statistically significantly improved for patients in the pemetrexed arm and median PFS appears greater for patients with adenocarcinoma histology compared with patients in the whole non-squamous group. Similar significant differences between the trial arms in respect of tumour response and disease control rates are noted. Survival rates at one year are substantially greater in the pemetrexed arm compared with the control arm. At two years, this difference is minimal.

Table 4-6 Key results of the JMEN trial for patients with adenocarcinoma histology

Endpoint	Pemetrexed (n = 222)	Placebo (n = 106)	HR (95% CI)	p value
Primary				
PFS (months) median	4.7	2.6	0.45 (0.35-0.39)	<0.00001
Secondary				
OS months (median)	16.8	11.5	0.73 (0.56-0.96)	0.026
Tumour response (%) (CR + PR)	8.1	2.8		0.090
Disease control rate (%) (CR+PR+SD)	61.0	33		<0.001
Survival rate at 1 year (%)	67	47		
Survival rate at 2 year (%)	29	26		

CI=confidence interval; CR=complete response; HR=hazard ratio; OS=overall survival; PFS=progression free survival; PR=partial response; SD= stable disease

Additional analyses requested by the ERG

The ERG requested that the manufacturer provide OS and PFS results by subgroup for the non-squamous population. The subgroups included disease stage (separating stage IIIb from stage IV), response status prior to maintenance therapy (with outcomes assessed as CR, PR or SD), first-line treatment (outcomes according to first-line regimen), first platinum treatment (cisplatin separately from carboplatin) and ECOG performance status (PS 0 separately from PS 1). The manufacturer's responses are shown in Table 10-3 and Table 10-4 of Appendix 3. In Table 10-3, on the outcome of PFS, females were shown to have a statistically significantly better outcome. For OS, females and East Asians (excluding those from the Indian subcontinent) appear to show better survival outcomes.

The results outlined in Table 10-4 should be viewed with caution as the trial was not powered to demonstrate statistical differences between the subgroups and the analyses were not adjusted for potential prognostic or confounding factors. However, the results indicate that for PFS there was a statistically significant effect for pemetrexed compared to placebo across all subgroups. For OS, the unadjusted HRs indicate a statistically significant effect of pemetrexed only for patients with stable disease following induction therapy, patients treated with paclitaxel or a taxane-based CTX, patients who received carboplatin as induction therapy and patients with PS 0. These trends may warrant further research in future trials.

4.2.3 Quality of life

The HRQoL outcome in the MS is time to worsening of symptoms, as measured by patient responses to the LCSS. Patients were asked to complete one questionnaire at baseline, at every cycle thereafter on approximately day 21, upon discontinuation from study, and approximately 30 days after the last dose of study drug. However, the manufacturer highlights (MS, pg45) that the rates of censoring/missing data were very high and so the statistical power and interpretation of findings are limited. In particular, it is noted in the Clinical Study Report (CSR, pg120) that the most commonly reported reason for not completing the LCSS was failure by the investigative site to administer the questionnaire due to concerns regarding patient welfare. The MS notes that this is a common problem in clinical trials, particularly within the study of cancer (MS, pg45). The ERG considers the lack of data as a serious limitation; HRQoL is an important Consideration for patients with lung cancer. In addition, maintenance therapy is a new area of research and may identify significant HRQoL aspects not highlighted at other points of treatment.

The summary of time to worsening of symptoms results is shown in Table 4-7. Responses for time to worsening of pain and haemoptysis were significantly different for patients in the pemetrexed arm compared to patients in the placebo arm; median time to worsening of symptoms of haemoptysis was not calculated because of high censoring/missing data rates. No statistically significant differences were noted between treatment arms on any other factor included in the scale; however lack of power due to poor data collection may have contributed to this result. Where incomplete figures for confidence intervals are shown, these are also incomplete in the MS and the CSR.

Table 4-7 Time to worsening of symptoms (non-squamous population)

Individual LCSS Score	**n**	**Pemetrexed median (months) (95% CI)**	**Placebo median (months) (95% CI)**	**HR (95% CI)[a]**	**p value[a]**
Loss of appetite	204	4.27 (5.78-)	4.63 (2.96-)	1.113 (0.81-1.52)	0.501
Fatigue	217	3.19 (2.79-6.28)	3.09 (2.43-3.98)	0.957 (0.71-1.28)	0.770
Cough	166	7.13 (4.73-)	6.44 (3.52-15.64)	0.883 (0.63-1.24)	0.471
Dyspnoea	179	10.71 (4.37-)	3.55 (2.79-15.61)	0.836 (0.61-1.15)	0.271
Haemoptysis[b]	33	-	15.61 (15.61-)	0.445 (0.22-0.90)	0.024
Pain	183	8.41 (5.16-12.45)	4.90 (2.79-15.61)	0.693 (0.51-0.95)	0.022
Symptom distress	209	4.50 (3.65-6.08)	3.68 (2.79-15.61)	0.879 (0.65-1.19)	0.403
Interference with activity level	182	7.82 (5.16-)	3.71 (2.43-15.61)	0.794 (0.58-1.09)	0.152
Global quality of life	188	7.20 (4.53-15.90)	3.68 (2.79-6.28)	0.795 (0.58-1.09)	0.149

CI= confidence interval; HR= hazard ratio; LCSS = Lung Cancer Symptom Scale; n= number patients with symptom
a unadjusted HR and p-value from Cox model with treatment as only co-factor
b median time to worsening of symptoms not calculated due to high level of censoring/missing data

4.2.4 Safety

Table 4-8 shows patient exposure to treatment. According to the MS (MS, pg51), there were no statistically significant differences in the incidence of drug-related grade 3 or 4 toxicities between patients who received ≤6 cycles of pemetrexed and those who received >6 cycles.

The MS further reports that AEs experienced by the non-squamous population were consistent with those of the overall trial population that included patients with squamous histology. A summary of AEs is provided in Table 4-9 which shows differences between treatment arms for anaemia,

neutropenia and fatigue which may be clinically important, albeit not always achieving statistical significance.

Rates of transfusions and supportive care are noted in the MS (MS, pg53) although it is unclear whether the figures relate to the total trial population, which includes patients with squamous cell cancer, or the non-squamous population alone. The results summarised in Table 4-10 were taken from the manufacturer's response to the ERG's clarification request. Differences between the trial arms are noted for rates of patients who received at least one transfusion, patients with at least one hospitalisation and hospitalisation due to drug-related toxicity. As would be expected, there were significantly more patients in the pemetrexed arm who discontinued treatment due to AEs.

Table 4-8 Patient exposure to treatment in the JMEN trial (non-squamous population)

Number of cycles	**Pemetrexed (n = 326)**	**Placebo (n = 156)**
Median	6	3
Mean	8	4.5
Standard deviation	8.62	5.32
No (%) completing at least 6 cycles	175 (53.8)	39 (25.0)
No (%) completing at least 10 cycles	82 (25.2)	11 (7.1)

Table 4-9 Percentage of patients with grade 3 or 4 toxicities (non-squamous population)

Grade 3 or 4 toxicity	**Pemetrexed (%) (n = 326)**	**Placebo (%) (n = 156)**	**p-value**
Anaemia	2.5	0	0.058
Neutropenia	2.8	0	0.035
Fatigue	3.7	0.6	0.070
Nausea	0.6	0.6	1.000
Vomiting	0.3	0	1.000

Table 4-10 Transfusions and other supportive care rates (non-squamous population)

Event	**Pemetrexed (n=325) (%)**	**Placebo (n=156) (%)**	**p-value**
Transfusion	8.9	2	NS
Patients with at least one hospitalisation	15.7	12.8	0.493
Hospitalisation due to drug-related AEs	5.2	0	<0.001
Patients discontinuing treatment due to AEs	4.8	1.4	0.027

AE= adverse event; NS= not stated

4.3 Summary of submitted evidence

4.3.1 Clinical results

- Results of the JMEN trial[2] demonstrate that in the patients with non-squamous histology, pemetrexed + BSC as maintenance therapy significantly increases median PFS when compared to placebo + BSC (4.5 vs. 2.6 months; HR 0.44; 95% CI 0.36-0.55, p<0.00001). Likewise, there was a statistically significant increase in median OS in the pemetrexed + BSC group when compared to the placebo + BSC group (15.5 months vs. 10.3 months; HR 0.70; 96% CI 0.56-0.88, p=0.002). Tumour response rate was statistically significantly greater in patients treated with pemetrexed compared to the control arm (7.4% vs. 1.9% p=0.018) as was disease control rate (57.7% vs. 32.7% p<0.001). Survival rates at one year were substantially greater in the pemetrexed arm compared to the control arm (60% vs. 42%). These differences were smaller at two years (28% v 22%).

- In the subgroup of patients with adenocarcinoma, the difference in median PFS was statistically greater in the pemetrexed + BSC arm when compared to the control arm (4.7 months vs. 2.6 months; HR 0.45; 95% CI 0.35-0.59, p<0.00001). Likewise, there was a statistically significant increase in OS in the pemetrexed-treated patients when compared to controls (16.8 months vs. 11.5 months; HR 0.73; 96% CI 0.56-0.96, p=0.026). Tumour response rate was statistically significantly greater in patients in the pemetrexed arm (8.1%) compared to the control arm (2.8%) as was the disease control rate (61.0% vs. 33% p<0.001). Survival rates at one year were substantially greater in the pemetrexed arm compared to the control arm (67% vs. 47%). These differences were smaller at two years (29% vs. 26%).

- The evidence for HRQoL was limited due to a high degree of censoring/missing data

- Patients treated with pemetrexed had statistically significantly higher rates of grade 3 or 4 neutropenia, experienced higher rates of transfusions, and hospitalisation due to drug toxicity. Patients treated with pemetrexed were more likely to discontinue treatment due to AEs.

4.3.2 Clinical issues and uncertainties

- The clinical results are based on a single trial which was generally well-designed but had a number of flaws in its execution and reporting. In particular:
    - the primary endpoint of the trial OS was changed to PFS during the trial without clear justification,
    - the key clinical evidence is derived from a histological subgroup of the RCT population. This subgroup was not included in the stratification of the randomisation procedure and the trial was not powered to perform this subgroup analysis.
- The generalisability of the trial to UK practice is also uncertain, due to the trial:
    - having no UK centres in the trial,
    - having a high proportion of Asian patients which are known to do better than other ethnicities,
    - having a high proportion of second-line treatments than are not commonly prescribed in the UK, which may affect OS and PFS,
    - excluding patients treated with first-line vinorelbine or pemetrexed, both of which are available in the UK,
    - having unlimited cycles of pemetrexed, which is unlikely to occur in the UK.
- The HRQoL data presented were also very limited, despite the trial being designed to collect it. This means it is very difficult to determine how patients' HRQoL will be affected by pemetrexed in a maintenance setting.
- The ERG is uncertain that the trial results give a true reflection of the OS, PFS and HRQoL benefits that could be expected in UK clinical practice.

5 ECONOMIC EVALUATION

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by the manufacturer of pemetrexed. The two key components of the economic evidence presented in the MS are (i) a systematic review of the relevant literature, and (ii) a report of the manufacturer's *de novo* economic evaluation. See Table 5-1 for a summary of key information points. The manufacturer also provided an electronic version of the Excel-based economic model.

Table 5-1 Key information in the MS

Key information	**Pages in the MS**	**Key tables/figures in the MS**
Details of the systematic review of the economic literature	57-58	Figure 7
Technology, patients, comparator, perspective and time horizon	59-65	
Framework for model-based evaluation	65-83	Tables 15-20 Figures 9-15
Clinical evidence used in economic evaluation	84-85	
Measurement and valuation of health benefits	86-91	Tables 21-22
Resource identification, measurement and valuation	91-102	Tables 23-34
Methods of sensitivity analysis and validity assessment	102-104	Tables 35
Results – base case analysis	105-106	Tables 36-38
Results – subgroup analysis	106-108	Tables 39-41
Results – sensitivity analysis	108-116	Tables 42-45
Results – end of life criteria	118-122	Tables 46-48
Assessment of factors relevant to the NHS and other parties	123-130	Tables 49-56

5.2 Overview of manufacturer's cost-effectiveness review

The manufacturer conducted a review of the published literature to retrieve cost-effectiveness studies of CTX agents used in the maintenance treatment of lung cancer. So as to be as comprehensive as possible, the manufacturer also specifically searched for first-line therapy, with the forethought that studies of first-line therapy may also include a component of maintenance therapy, which seems appropriate given the novelty of maintenance treatment.

5.2.1 Identification and description of studies

The MS appendices included full details of the electronic search strategy and date and language limits. The databases searched included EMBASE, MEDLINE, NHS EED and HEED, and the manufacturer also stated that hand searching of reference lists of retrieved articles was also undertaken. As mentioned above, the search was broad enough to include both maintenance and first-line therapies.

The manufacturer's inclusion criteria were limited to NSCLC and publications reporting both costs and benefits of any of the following interventions: gemcitabine, pemetrexed, docetaxel, vinorelbine, paclitaxel, erlotinib, bevacizumab, cetuximab, gefitinib, BSC or placebo. The exclusion criteria barred studies which were not full cost-effectiveness studies, were cost-minimisation studies, were conducted in the wrong population, were second-line therapy, had the wrong intervention, or were a review or a duplicate publication.

The manufacturer's search strategy identified 120 possibly relevant articles. Subsequently, 17 were retrieved for detailed evaluation and a further two studies were identified via hand searching activities. Of these 19 studies, five were excluded after application of the inclusion/exclusion criteria and the remaining 14 were first-line therapy only and hence excluded on the grounds that they were not directly related to the decision problem.

5.2.2 Summary and conclusions

The manufacturer's review of the published cost-effectiveness literature describing maintenance therapy for patients with NSCLC did not identify any relevant cost-effectiveness studies. The ERG is satisfied with the manufacturer's search strategy and is reasonably confident that the manufacturer did not miss any relevant published articles. However, the manufacturer did not appear to undertake any searches of the unpublished literature, which may mean that relevant unpublished studies were omitted. In summary, the ERG believes that the manufacturer's systematic review of the literature addressed the decision problem in a clear and transparent manner, and that the likelihood that the manufacturer missed relevant published studies is minimal.

5.3 Overview of manufacturer's economic evaluation

The manufacturer undertook a *de novo* economic evaluation of pemetrexed for the maintenance treatment of advanced or metastatic non-squamous NSCLC in patients who had not progressed following four cycles of first-line CTX with a platinum doublet containing gemcitabine, paclitaxel or docetaxel only. A schema of the manufacturer's economic evaluation is provided in Figure 5-1.

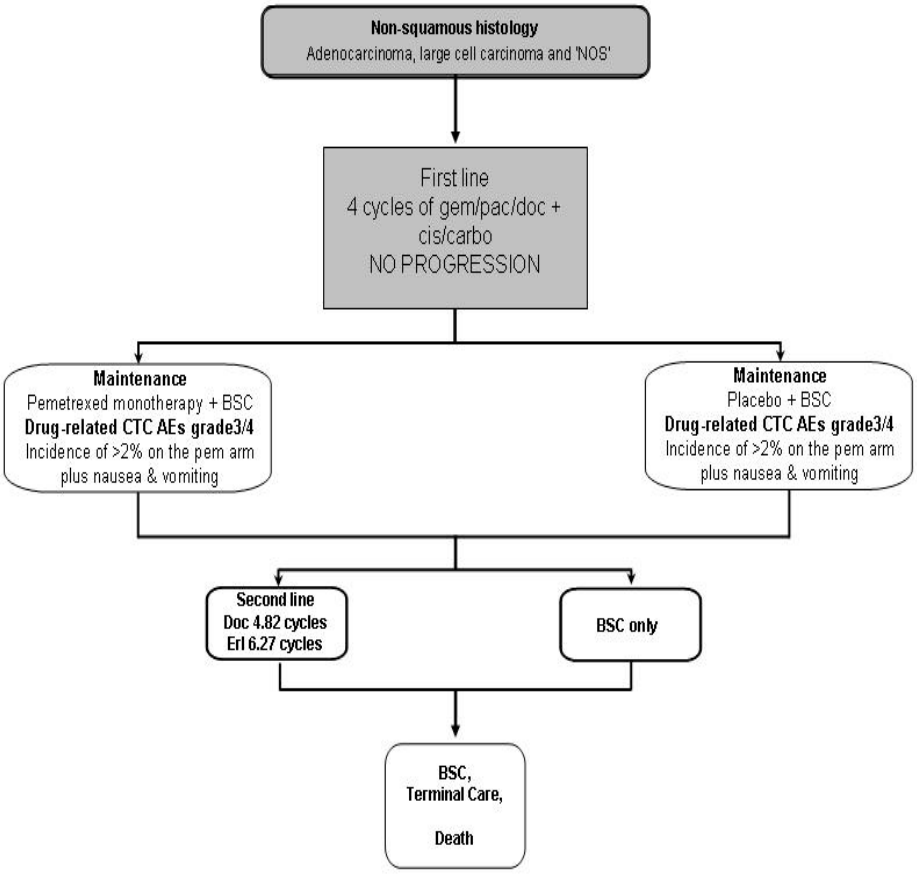

Figure 5-1 Schema of manufacturer's economic evaluation

5.3.1 Description of manufacturer's economic model

The manufacturer presents a simple trial-based model with an extrapolation component which takes the 29 month trial data out to six years. The extrapolation is based on the exponential survival function, (see Section 5.3.3 below for further details) and is only applied to the post-trial period as the first 29 months are taken directly from the trial. It is worth noting that AEs are not captured in the model, but are explored in the sensitivity analysis (SA).

Patients enter the model at the start of maintenance therapy, which is assumed to start immediately after four cycles of first-line CTX in patients who have not progressed, i.e. patients who have had a response (full or partial), or patients with stable disease. The model has two arms: placebo and pemetrexed. Patients in the placebo arm are assumed to be receiving 'watch and wait' treatment and to be in receipt of BSC as needed. Patients in the pemetrexed arm receive pemetrexed every three weeks together with BSC as needed. However, in contrast to the JMEN trial[2] where patients in the active arm were given pemetrexed until disease progression (one patient received 55 cycles), the model imposes a capping rule which limits the maximum number of cycles of pemetrexed that can be given to patients. In the base case analysis this upper limit is set at 17 cycles. In the economic model this capping rule is only applied to the calculation of costs and does not impact upon the health outcomes, so that patients are assigned the full trial benefits of pemetrexed but with reduced costs. The manufacturer varies this capping rule in the SA; see section 5.3.11 for further details.

After patients progress in the model, they are eligible for second-line CTX. The rates of second-line CTX are taken directly from the JMEN trial:[2] 67% for placebo patients and 53% for pemetrexed patients. It is worth noting that 18.5% of placebo patients in the trial crossed over to pemetrexed after they had progressed (at which point patient status becomes unblinded). This may have inflated the rate of second-line CTX administered in the placebo arm, giving rise to an artificial differential in the rates of CTX between the two arms with the possibility of consequent unpredictable alterations to patient outcomes compromising the decision comparison. Second-line CTX is composed of either docetaxel or erlotinib monotherapy which are assumed to have the same unit costs and the same efficacy (which is captured in the trial and hence not specifically entered as a model parameter) but differential numbers of cycles. For patients not receiving second-line CTX, they continue to receive BSC alone.

Following second-line CTX, patients enter a terminal phase when they receive BSC only. The final 3-week period of life is designated as 'terminal care' to which a higher cost is assigned. The model

continues for a maximum of six years at which time point 99% of placebo patients and 96% of pemetrexed patients are expected to have died.

5.3.2 Parameters and values

The base case model parameters are shown in Table 5-2.

Table 5-2 Parameters in the manufacturer's model

Model inputs	Values	Source/Description
UTILITY		
Utility for pts treated with pem maintenance	0.66	Adapted from Nafees et al, 2008.[15] Corresponding to Not Progressed, average of Stable and Responding with no AE
Utility for pts treated with placebo maintenance	0.58	From Nafees et al, 2008.[15] Corresponding to "Stable with fatigue"
Utility for BSC post-progression (all years for placebo arm and 2nd year onwards for pem)	0.53	From Berthelot et al, 2000.[16] Corresponding to BSC post-progression (i.e. second-line phase)
Utility for BSC post-progression (pem arm, 1st year only)	0.54	Adapted from Berthelot, 2000.[16] Corresponding to BSC post-progression plus 0.01 increment due to better pain control with pem therapy
Utility for second-line CTX (both arms)	0.58	From Nafees et al, 2008.[15] Corresponding to Stable with Fatigue
Utility for terminal cycle (both arms)	0.47	From Nafees et al, 2008.[15] Corresponding to Progression
COSTS		
Pemetrexed (100mg vial)	£160.00	100mg vial MIMS June 2009[17]
Pemetrexed (500mg vial)	£800.00	500mg vial MIMS June 2009[17]
Pemetrexed dose	500 mg/m^2	SPC dose
Docetaxel (20mg vial)	£162.75	MIMS June 2009[17]
Docetaxel (80mg vial)	£534.75	MIMS June 2009[17]
Docetaxel dose	75 mg/m^2	SPC dose
Erlotinib (150 mg, 30 tablet pack)	£1,394.96	Cost per 150mg tab £1631.53/30 tabs (MIMS June 2009)[17] + 14.5% discount from list
Erlotinib (per dose)	£46.50	150mg per day as per SPC
Cost of outpatient administration of pemetrexed	£153.00	National Schedule of Reference Costs 2007-08 CTX delivery outpatients SB12Z[18]
Cost of outpatient administration of docetaxel	£208.00	National Schedule of Reference Costs 2007-08 CTX delivery outpatients SB14Z[18]
Cost of outpatient administration of erlotinib	£125.25	National Schedule of Reference Costs 2007-08 Oral CTX delivery outpatients SB11Z[18]
Average cost of BSC per cycle (no active chemo)	£66.36	Adapted from NICE palliative care document 2004[19] inflated to 2008 costs
Average cost of BSC per cycle (active chemo)	£33.18	Adapted from NICE palliative care document 2004 [19]inflated to 2008 costs (assume no radiotherapy during active chemo)
Average cost of terminal care	£2,588.25	Adapted from NICE palliative care document 2004[19] inflated to 2008 prices. One-off cost applied in the last cycle of life
ASSUMPTIONS/INPUTS		
Mean cycles of pemetrexed maintenance therapy	5.84	From JMEN trial[2] for non-squamous patients, assuming max of 17 cycles
Mean cycles of placebo/BSC as maintenance therapy	4.50	From JMEN trial[2] non-squamous population
Mean number of cycles of second-line chemo (docetaxel)	4.82	NICE TA162 - ERG report document: Erlotinib for the treatment of relapsed NSCLC, page 57[20]

Mean number of cycles of second-line chemo (erlotinib)	6.27	NICE TA162 - ERG report document: Erlotinib for the treatment of relapsed NSCLC, page 57[20]
Use of second-line chemo for pemetrexed maintenance patients	53.2%	from JMEN trial[2] (non-squamous patients)
Use of second-line chemo for placebo maintenance patients	67.3%	from JMEN trial[2] (non-squamous patients)
% patients receiving erlotinib as second-line chemo	27.3%	Lilly market share data (reports chemotherapies with market share >5% only)
% patients receiving docetaxel as second-line chemo	73.7%	Lilly market share data (reports chemotherapies with market share >5% only)

JMEN= JMEN clinical trial; SPC=summary of product characteristics; BSC=best supportive care; pem=pemetrexed; CTX=chemotherapy

5.3.3 Treatment effectiveness within the MS

The clinical data used in the manufacturer's economic evaluation are taken directly from the JMEN trial,[2] which is described in detail in section 4 of this report.

As the JMEN trial[2] OS data are censored at 29 months when 30% of trial patients are still alive, the manufacturer chose to extrapolate the trial OS data for 43 months to give a time horizon of six years. The manufacturer went through a detailed process of curve-fitting to generate curves with the 'best fit'. In the base case analysis, an exponential hazard function was utilised to project outcomes from the end of the trial period to six years. Weibull exponential hazard functions were explored in a supplementary analysis (see MS page 66-72 for more details).

5.3.4 Population

The population in the manufacturer's economic evaluation is based on the JMEN trial[2] population. That is stage IIIb/IV patients who have received four cycles of first-line CTX (based on a platinum doublet including gemcitabine, docetaxel or paclitaxel only) and whose disease has not progressed. As pemetrexed is only licensed for patients with non-squamous histology, the manufacturer's economic evaluation excludes JMEN trial[2] patients with squamous disease.

As the largest patient subgroup in the JMEN trial[2] was patients with adenocarcinoma, the manufacturer presents results for this subgroup of patients. No other subgroups were presented.

5.3.5 Comparator technology

In the JMEN trial[2] the comparator was a placebo saline infusion (to maintain blinding) together with BSC as needed. In the economic evaluation, however, the comparator is often still called placebo but in reality the manufacturer appears to mean 'watch and wait' together with BSC as needed, which is currently the only available treatment alternative to maintenance therapy for patients with NSCLC.

5.3.6 Health related quality of life

Health related QoL data were not available from the JMEN trial.[2] Instead, the manufacturer undertook a systematic review of the literature to identify relevant HRQoL data for use in the economic evaluation. Fifteen studies were identified by the manufacturer but in the end a single study funded by the manufacturer (Nafees, 2008[15]) was utilised as a source for the majority of the utility values employed in the model, and supplemented as needed with values taken from one other study (Berthelot, 2000[16]). Nafees was a Lilly sponsored UK study conducted in 100 members of the general public using standard gamble interview techniques; whilst Berthelot was a Canadian study conducted in NSCLC patients using visual analogue scales. See Table 5-2 above for a list of utility values used in the MS.

5.3.7 Resources and costs

Quantities of the main resource items included in the economic evaluation such as medication, CTX administration and occurrence of AEs were taken directly from the JMEN trial.[2] See Table 5-3 for a summary of the resource and unit cost sources used in the MS. Resource use for BSC and terminal care were taken from the published literature. Unit cost data were taken from Monthly Index of Medical Specialities (MIMS),[17] NHS reference costs,[18] the published literature[21,22] and surveys of clinical experts. All unit costs were inflated (as necessary) to the price year 2008, apart from MIMS,[17] which was available as 2009 data. For actual cost data used in the model see Table 5-2 above and pages 94-99 in the MS.

5.3.8 Perspective, time horizon and discounting

Costs are estimated from the perspective of the NHS, and outcomes are expressed as QALYs; both of which were captured over a six-year time horizon (which is assumed to be a life-time horizon). Costs and outcomes were discounted at a rate of 3.5%, in line with current NICE guidance.{NICE, 2008 #54}

5.3.9 Model validation

To validate the model the manufacturer compared OS figures from the model with the OS figures from the JMEN trial,[2] (i.e. internal validation) see page 104 of the MS. As the model was trial-based the manufacturer did not believe that any other validation (e.g. external validation) was necessary.

Table 5-3 Resource use and unit cost data sources used in the MS

Resource	Utilisation rate data source	Unit cost data source
Medication		
CTX acquisition	JMEN trial[2] & SPC	MIMS[17] July (2009)[17]
Concomitant medication	SPC	
Administration		
CTX administration (including concomitant medication)	JMEN trial[2] & SPC	NHS reference costs 2007-2008[18]
Adverse events		
Neutropenia	JMEN trial[2]	Survey of clinical experts Duran et al 2008,[21] Hanna 2004[22]
Fatigue	JMEN trial[2]	Duran et al 2008,[21] Hanna 2004[22]
Nausea/vomiting	JMEN trial[2]	Duran et al 2008,[21] Hanna 2004[22]
Anaemia	JMEN trial[2]	Duran et al 2008,[21] Hanna 2004[22]
Best supportive care	Literature review (see Table 34 and Appendix 10 of MS).	Literature review (Table 34 and Appendix 10 of MS) and NICE palliative care document 2004[19]
Terminal care	NICE palliative care document 2004[19]	Literature review (see Table 34 and Appendix 10 of MS) and NICE palliative care document 2004[19]

JMEN= JMEN clinical trial; SPC=summary of product characteristics; MS=manufacturer's submission; CTX=chemotherapy

5.3.10 Results included in the MS

Base case results

The base case results (non-squamous population) generated by the manufacturer's model are presented below in Table 5-4. The incremental cost-effectiveness ratio (ICER) for the non-squamous population is £33,732 per QALY, which is based upon use of the exponential survival function.

Table 5-4 Base case results

	Pemetrexed (Pemetrexed / BSC)	**Placebo (Watch and wait / BSC)**	**Incremental**
Cost results			
Maintenance therapy plus administration	£9,903	£299	£9,605
Second-line therapy plus administration	£3,570	£4,516	-£946
AE cost	£34	£5	£29
BSC (with CTX)	£105	£133	-£28
BSC (without CTX)	£1,329	£847	£481
Terminal care	£2,514	£2,518	-£4
Total costs	**£17,455**	**£8,318**	**£9,137**
Effectiveness results			
Total LYG	1.7	1.26	0.44
Total QALYs	0.97	0.70	**0.27**
ICER			
Cost per LYG			£20,562
Cost per QALY			**£33,732**

BSC= best supportive care; AE=Adverse event; QALY=quality adjusted life year; LYG=life year gained; ICER=incremental cost-effectiveness ratio; CTX=chemotherapy

NB: when the manufacturer uses the Weibull function the ICER increases to £36,386 per QALY (see MS, page 109).

Subgroup results

The results for the adenocarcinoma subgroup are presented below in Table 5-5. The ICER for the adenocarcinoma population is £39,364 per QALY, which is based upon use of the exponential survival function.

Table 5-5 Adenocarcinoma subgroup results

	Pemetrexed (Pemetrexed / BSC)	**Placebo (Watch and wait / BSC)**	**Incremental**
Cost results			
Maintenance therapy plus administration	£10,446	£305	£10,141
Second-line therapy plus administration	£3,679	£4,654	-£975
AE cost	£22	£1	£21
BSC (with CTX)	£71	£109	-£37
BSC (without CTX)	£1,481	£1,072	£409
Terminal care	£2,429	£2,432	-£3
TOTAL COSTS	**£18,129**	**£8,574**	**£9,554**
Effectiveness results			
Total LYG	1.87	1.45	0.42
TOTAL QALYS	1.03	0.79	**0.24**
ICER			
Cost per LYG			£22,788
COST PER QALY			**£39,364**

BSC= best supportive care; AE=Adverse event; QALY=quality adjusted life year; LYG=life year gained; ICER=incremental cost-effectiveness ratio; CTX=chemotherapy

NB: when the manufacturer uses the Weibull function the ICER increases to £42,922 per QALY (see MS, page 109).

5.3.11 Sensitivity analyses

The manufacturer undertook six scenario analyses and 36 one-way sensitivity analyses (SA) (see the MS page 108-116 for the full details). The scenario analyses had the biggest impact upon the size of the ICER producing ICERs ranging from £14,823 to £134,666 per QALY for the non-squamous population, see Table 5-6 below. However, the extreme outliers (scenario 3 and 4) were apparently produced without including uncertainty in the parameterised survival projections and so may not be representative of the range of possible ICERs.

The majority of the one-way SA (35/36) produced ICERs in the range of £20,000 to £50,000 per QALY, with most being very close to the base case ICER of £33,732 per QALY (see Table 45 in the MS). However, one of the analyses presented gave an ICER of £105,826 per QALY, based on a lower incremental survival value of 1.15 months (base case 5.5 months) for the non-squamous population.

The ERG notes that probabilistic sensitivity analysis (PSA) was not undertaken by the manufacturer, which is a major limitation of the MS given (i) the closeness of the submitted ICER to the NICE threshold range and (ii) the numerous adjustments used in the model to account for UK practice or to simplify the decision problem. Furthermore, the manufacturer does not provide sufficient justification for the choice of parameters and parameter values varied in the one-way SA performed; often parameter variation appears to be arbitrary in nature and not driven by clinical advice or alternative data sources. The manufacturer also claims that the structure of the extrapolation model does not lend itself to undertaking a PSA, but this is not a supportable position as standard errors and correlations for exponential and Weibull functions are readily available from statistical analysis packages.

The ERG is therefore of the opinion that it is not clear that the true uncertainty surrounding the decision problem has been fully explored by the manufacturer.

Table 5-6 Scenario analysis for the non-squamous population

Scenario	Incremental QALY	Incremental cost	ICER
Scenario 1 (exponential distribution): Mean number of cycles as per JMEN (8 cycles for non-squamous population) BSA = 1.82 m^2 Per vial costing AE disutility applied to PEM	0.2847	£13,379	£46,992
Scenario 1 (Weibull distribution): Mean number of cycles as per JMEN (8 cycles) BSA = 1.82 m^2 Per vial costing AE disutility applied to PEM	0.2637	£13,334	£50,564
Scenario 2 (exponential distribution): Cycles capped at 10 (equivalent to mean of 4.61 cycles for non-squamous population) BSA = 1.8m^2 Per mg costing Pain benefit in second-line (Doyle) No AE disutility applied to PEM	0.2966	£6,813	£22,972
Scenario 2 (Weibull distribution): Cycles capped at 10 (equivalent to mean of 4.61 cycles for non-squamous population) BSA = 1.8m^2 Per mg costing Pain benefit in second-line (Doyle)	0.2756	£6,767	£24,558
Scenario 3 (conservative efficacy): Mean number of cycles as per JMEN (8 cycles for non-squamous population) BSA = 1.82 m^2 Per vial costing Efficacy (lower 95%CI for Pem & upper 95%CI for BSC) AE disutility applied to PEM	0.0963	£12,970	£134,666
Scenario 4 (optimistic efficacy): Cycles capped at 10 (equivalent to mean of 4.61 cycles for non-squamous population) BSA = 1.8m^2 Per mg costing Pain benefit in second-line (Doyle) Efficacy (upper 95%CI for Pem & lower 95%CI for BSC) No AE disutility applied to PEM	0.4876	£7,227	£14,823

BSA= body surface area; AE=adverse event; PEM=pemetrexed; QoL=quality of life; CI=confidence interval; BSC=best supportive care; JMEN=JMEN clinical trial

5.4 Assessment of the manufacturer's economic model

Table 5-7 shows how closely the manufacturer's submitted economic evaluation accords with the requirements for a base case analysis as set out in the NICE reference case checklist.[23] It is clear that the manufacturer has attempted to adhere to the NICE reference case which is commendable, but unfortunately on a number of points the manufacturer has not succeeded. In particular, the manufacturer has not undertaken PSA which means that uncertainty may not have been fully accounted for, and although the manufacturer undertook discounting using the appropriate 3.5% rate, they did not apply the discount rate correctly (see section 5.5 below for more details). Furthermore, the source of utility values used in the economic model may not be appropriate to the decision problem.

Table 5-8 summarises the ERG's appraisal of the economic evaluation conducted by the manufacturer using the Drummond 10-point checklist.[24] The manufacturer's model fails on a number of issues, crucially the valuing of costs and benefits, and the assessment of the uncertainty surrounding the model results (see section 5.5 below for more details).

Table 5-7 NICE reference case checklist

Attribute	Reference case	Does the *de novo* economic evaluation match the reference case?
Comparator(s)	BSC	Watch and wait + BSC, which is the appropriate comparator
Perspective costs	NHS and Personal Social Services (PSS)	The economic evaluation is carried out from the perspective of the NHS. No PSS costs are described in the MS
Perspective benefits	All health effects on individuals	Health effects to the individual are captured via QALYs
Form of economic evaluation	Cost-effectiveness analysis	Cost-effectiveness analysis
Time horizon	Sufficient to capture differences in costs and outcomes	The time horizon chosen was a lifetime horizon, which for this patient group was believed to be within six years. This appears appropriate
Synthesis of evidence on outcomes	Systematic review	All survival data are derived (and extrapolated) from the JMEN RCT the only relevant clinical trial identified by systematic review
Outcome measure	Quality adjusted life years (QALYs)	QALYs were used, which is appropriate
Health states for QALY	Described using a standardised and validated instrument	Quality of life data were not available from the JMEN trial,[2] therefore two published QoL studies were utilised, primarily Nafees (standard gamble in general public),[15] with data from Berthelot (VAS in NSCLC patients)[16] as required. This is not ideal and is further hampered by the selective use of the utility values presented in the studies. Furthermore, the QoL study was not specifically designed to capture the QoL of patients on maintenance therapy
Benefit valuation	Time-trade off or standard gamble	The main QoL[15] study utilised standard gamble interview techniques, which is acceptable
Source of preference data for valuation of changes in HRQL	Representative sample of the public	The main QoL study[15] was based on responses from 100 members of the general public. It is not clear how representative this sample is of the UK adult population.
Discount rate	An annual rate of 3.5% on both costs and health effects	Benefits and costs have been discounted using a rate of 3.5% after year one, but this rate was not applied correctly
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	All QALYs estimated by the economic model have the same weight
Sensitivity analysis	Probabilistic sensitivity analysis (PSA)	PSA was not undertaken by the manufacturer, and the manufacturer's handling of sensitivity analysis was limited and not fully justified, which means uncertainty in the model may not have been fully accounted for

PSS= Personal Social Services; MS=manufacturer submission; RCT=randomised controlled trial; QoL=quality of life; QALYs=quality adjusted life years; PSA=probabilistic sensitivity analysis; ERG=Evidence Review Group

Table 5-8 Critical appraisal checklist

Item	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	The manufacturer answered the decision problem set by NICE
Was a comprehensive description of the competing alternatives given?	Yes	The manufacturer described the chosen comparator adequately
Was the effectiveness of the programme or services established?	Partially	The effectiveness of maintenance therapy is established using the JMEN RCT.[2] This trial, however, may not be representative of UK clinical practice as there were no UK centres. Furthermore, the trial did not allow patients to receive first-line vinorelbine or pemetrexed – both of which are licensed for use in the UK. Similarly, the trial allowed patients to receive pemetrexed and gefitinib as second-line therapy, which is not standard practice in the UK
Were all the important and relevant costs and consequences for each alternative identified?	Yes	The key costs and outcomes were identified
Were costs and consequences measured accurately in appropriate physical units?	Not consistently	For example, the BSA value used to calculate CTX costs does not represent NSCLC patients in the UK
Were the cost and consequences valued credibly?	No	Overall survival was not adequately modelled. Not all of the costs were adequately valued, for example, the extra testing required whilst on maintenance therapy may not have been fully accounted for in the model
Were costs and consequences adjusted for differential timing?	Partially	Costs and outcomes were discounted after 1 year, but the method of discounting was not applied correctly
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	ICERs (cost per QALY gained and cost per LYG) were presented for the base case population and the adenocarcinoma population
Was allowance made for uncertainty in the estimates of costs and consequences?	No	No PSA was undertaken. Univariate SA and scenario analysis were undertaken by the manufacturer but the choice of sensitivity analyses presented and the values chosen were not fully justified and do not always appear logical
Did the presentation and discussion of study results include all issues of concern to users?	Yes	The results are presented and discussed in detail and an end of life treatments case has been proposed by the manufacturer

ERG= Evidence Review Group; MS = manufacturer submission; QALY=quality adjusted life year; LYG=life year gained; SA=sensitivity analysis; PSA=probabilistic sensitivity analysis; ICER=incremental cost-effectiveness ratio

5.5 Detailed critique of manufacturer's economic model

5.5.1 Structure and assumptions

The submitted decision model is implemented as a series of Excel worksheets. The primary calculations are carried out on two worksheets which combine OS data obtained from Kaplan-Meier analysis of the JMEN trial[2] data, with parametric survival projection from the end of the trial data to the time horizon (six years). All other calculations are performed on a series of accounting tables, driven by results calculated on the survival worksheets.

The layout of the model is generally clear and the tables are clearly labelled. An 'Inputs' worksheet includes most of the important parameter values with a brief indication of the derivation of each item.

It should be noted that the ERG has concentrated attention on the model for the non-squamous population, since this relates to the licensed indication and the results for the smaller adenocarcinoma population are quite similar. In addition, the ERG have used the version of the model which uses the exponential (rather than Weibull) projection as the basis for comparison, this being the manufacturer's base case which results in an ICER for use of pemetrexed in maintenance therapy of £33,732 per QALY gained.

5.5.2 Major errors and omissions identified by ERG

Survival estimation and projection

The manufacturer's model combines direct use of trial outcome results with parametric projection modelling to estimate the mean OS time per patient from randomisation to the six year time horizon of the analysis. No direct use is made in the model of the primary trial outcome (PFS), which instead is replaced by the duration of maintenance therapy as a proxy.

The charts provided in the MS, showing the two parametric models developed (using exponential and Weibull functions) plotted together with the Kaplan-Meier survival curves from the trial arms (Figure 5-2), reveal a common problem for drug trials in that there is poor correspondence between the parametric models and the source data, especially at the beginning and end periods of the trial.

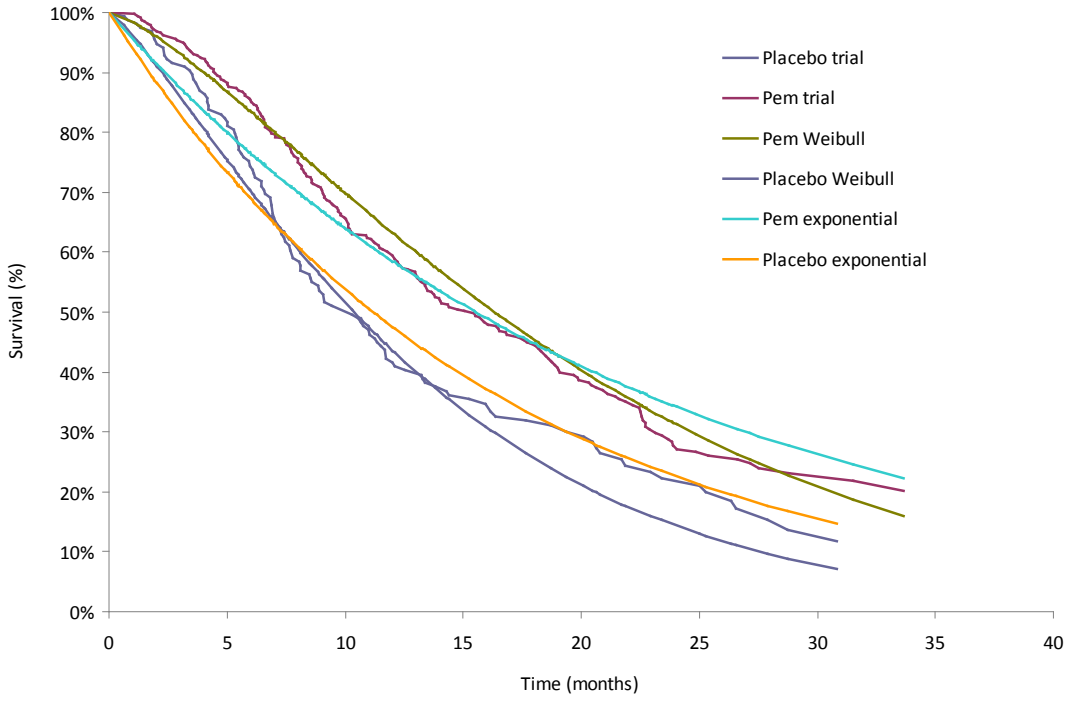

Figure 5-2 Overall survival: Kaplan-Meier survival plots from JMEN trial[2] data with fitted parametric models

There are several factors which may contribute to this lack of correspondence:

a) Trial inclusion/exclusion criteria normally include direct or indirect stipulations which minimise or remove altogether the likelihood of specific events occurring in the first few weeks of the trial. In this case, the JMEN trial[2] protocol requires that subjects should have "Estimated life expectancy of at least 12 weeks". This ensures that fewer deaths occur in the first four cycles of maintenance therapy than might normally be expected in a cohort of such patients.

b) The action of the new drug (pemetrexed in JMEN[2]) takes time to achieve its full effect, partly due to the pharmacokinetic profile of the drug, and partly due to the time required for the active agent to achieve its full effect at the target site(s). Conversely, when the period of active treatment comes to an end its effects are likely to dissipate gradually over several weeks. This mechanism may also be relevant at the start of treatment in this trial since there was no extended 'wash-out' period following initial CTX before maintenance therapy with pemetrexed commenced.

c) Additional confounding is potentially introduced by the availability of subsequent courses of active CTX which will further complicate the dynamic nature of the event HR following disease progression.

d) There is also the possibility that the patient population is essentially heterogeneous in relation to the event risk of interest, leading to progressive survivor bias as members of one subgroup suffers death at a faster rate than other patients.

As a consequence of these influences, it is not at all surprising that fitting a standard parametric survival function to the full clinical trial dataset rarely produces a satisfactory correspondence to the calculated survival trajectory. Moreover, since the reliability of fit at later periods is increasingly sensitive to diminishing patient numbers, calibrating a parametric function from the full patient data may be a particularly unsatisfactory basis for projecting events beyond the trial data collection period.

The model authors have correctly recognised that within the trial period the most reliable estimate of mean survival is obtained directly from the trial data, in the form of the area under the Kaplan-Meier survival curve. They then project beyond the trial up to the time horizon on the basis of two simple parametric models (exponential or Weibull) calibrated against the whole duration of the trial, despite the evident inaccuracy of these functions. The clearest evidence of the weakness of this approach is indicated by the discontinuities evident in the estimated survival plots when switching from the in-trial trend to the modelled trend (Figure 5-3).

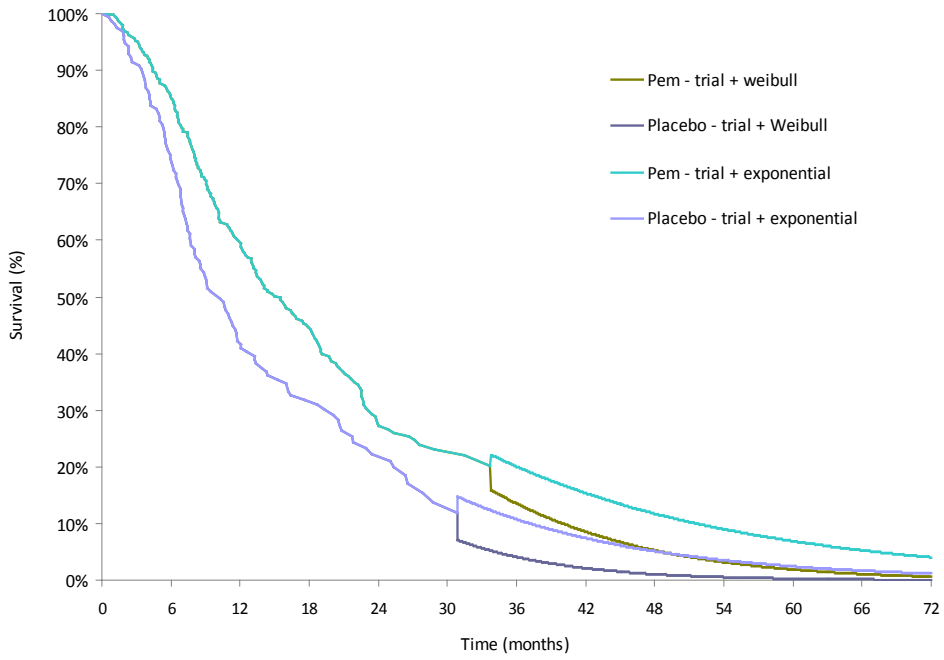

Figure 5-3 Trial survival plots combined with parametric projection functions

The model authors seek to overcome this problem by using a function derived from the cumulative hazard plot of the data, but constrained to commence at the last point on the Kaplan-Meier curve (Figure 5-4). This avoids the problem of a discontinuity, but cannot overcome the evident sudden change in slope occurring where the two parts of the survival function join, suggesting that this alternative may not be a satisfactory solution.

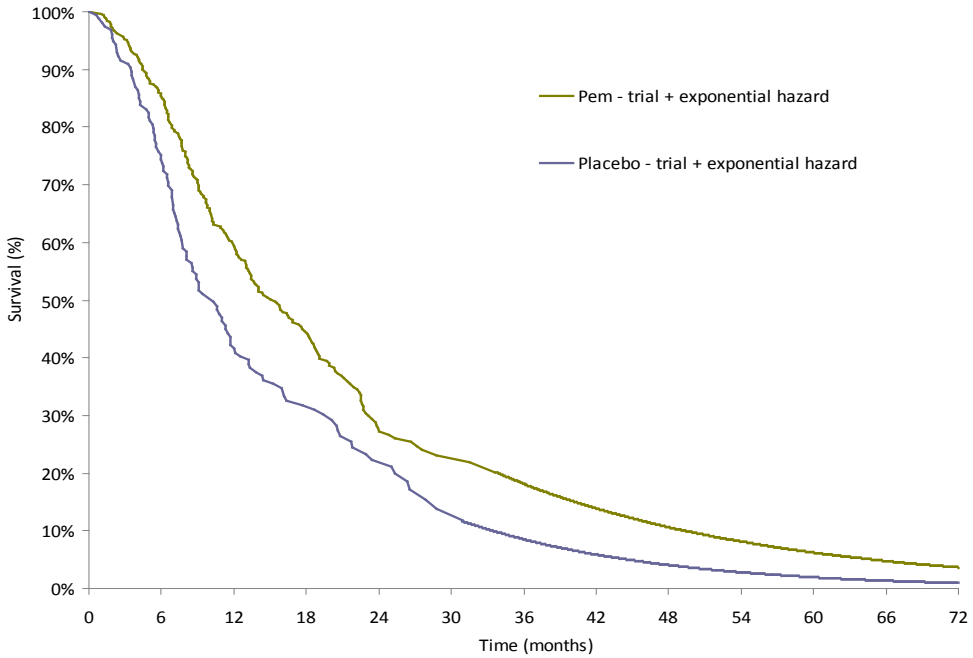

Figure 5-4 Trial survival plot conjoined with hazard-based projection function

A better approach is provided by an examination of the cumulative hazard plot over time. The objective should be to investigate whether a settled long-term pattern of hazard can be detected, which might be used as a more reliable basis for long-term projection when most of the transitory effects described above have resolved. In this case there appears to be a stable linear trend in the cumulative hazard apparent from eight months after randomisation which persists to the end of the trial, and which is present in both arms of the trial (Figure 5-5).

This suggests that the most reliable basis for projection would be to fit an exponential survival curve to the available patient data relating only to the period beyond eight months survival (where the risk dynamics have 'settled down'), and using this function for estimating survival beyond the trial period. This is likely to yield more accurate results in the latter parts of the study (Figure 5-6). This approach has been pursued by the ERG, making use of the extract of individual patient data from the JMEN trial[2] provided by the manufacturer in response to request in the clarification letter. The findings are described in detail below in Section 6 of this report.

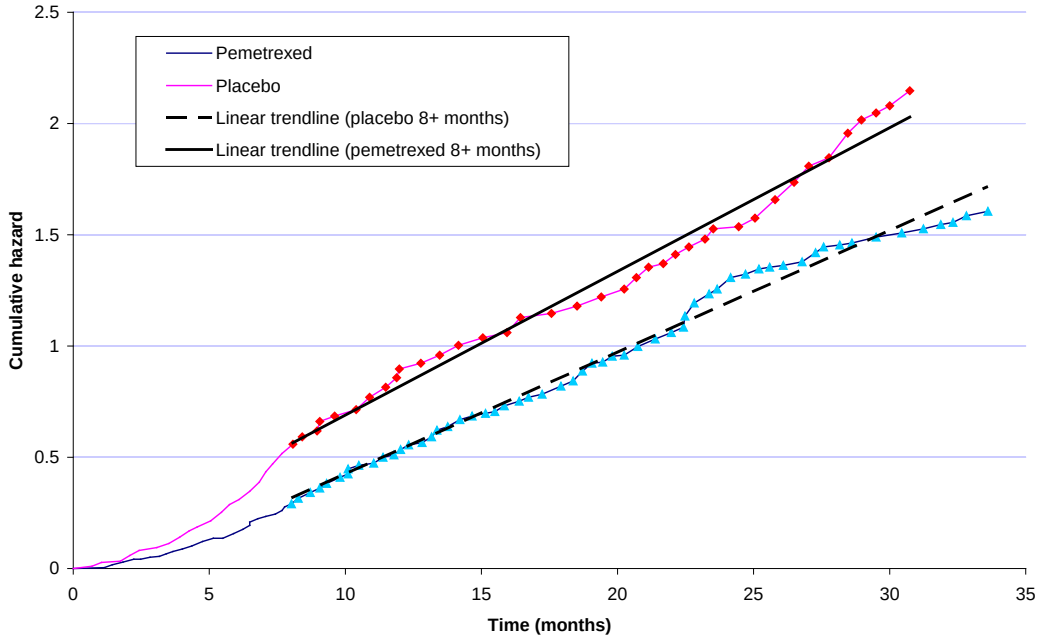

Figure 5-5 Linear hazard trendlines fitted to JMEN trial[2] data from 8 months onwards

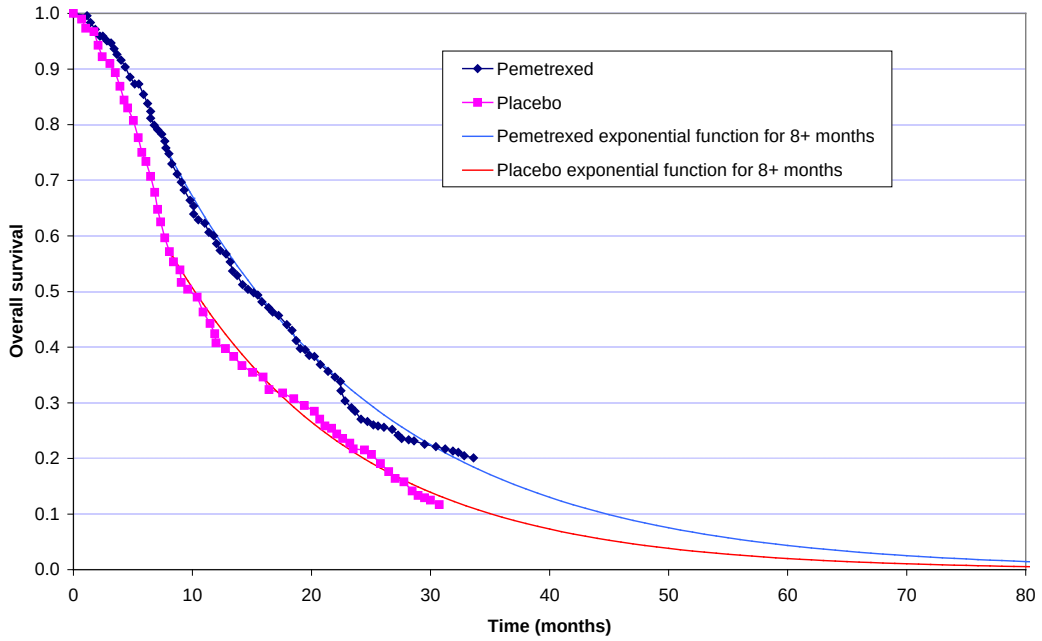

Figure 5-6 Overall survival exponential trendlines fitted to JMEN data for 8 months onwards

Base case for evaluation

The submitted base case scenario makes full use of the JMEN trial[2] results in most respects, with one important exception: the modellers have decided to limit the cost of maintenance therapy by imposing an arbitrary limit on the number of cycles of pemetrexed therapy that should be provided to any patient. The chosen limit (17 cycles) is far short of the recorded maximum seen in the trial population (55 cycles), and does not correspond with the manufacturer's own clinical advice which suggested 10 cycles. However, as there is no precedent for maintenance therapy being used for treating advanced NSCLC outside of clinical trials, it cannot be said that there is any relevant UK clinical practice on which to base such a judgement.

The importance of this modelling limitation is that it serves to constrain the costs of maintenance therapy (affecting 8.2% of pemetrexed patients and 15% of treatment cycles), but has no corresponding effect on the benefits accrued from use of pemetrexed, and therefore builds in an essential bias in the economic evaluation in favour of pemetrexed. There is no objective basis for assessing the impact of limiting the use of maintenance therapy on patient outcomes in the absence of objective evidence either from the JMEN trial[2] or from any other source. For this reason, the ERG considers that the most appropriate base case should include the full costs of maintenance therapy based on the cycles delivered during the JMEN trial,[2] combined with the full net benefits of maintenance therapy compared to the 'wait and see' strategy. Any limitation of drug use should be considered in a SA which should also explore possible correlations between reduced drug use and reduced outcome effects.

Utility values

Utility values in the submitted model are primarily drawn from a study by Nafees et al.[15] Although the methods used to sample and derive utility estimates might be questioned, the relevant values appear to be broadly consistent with those reported elsewhere for NSCLC patients. However, the manner in which values have been selected for use in the model appears to be arbitrary and unduly favours the pemetrexed arm of the evaluation. Although at randomisation it can be assumed that patients in both arms of the trial are on average in the same clinical condition (stable or responding to first-line CTX) and experience similar HRQoL, they have different utility values applied: 0.66 for pemetrexed patients, but only 0.58 for placebo patients. This is counterintuitive as referring back to the JMEN trial[2] (non-squamous population), the rate of grade 3/4 fatigue was noticeably higher in the pemetrexed arm (3.66%) than in the placebo arm (0.64%).

The explanation is that the utility value used for pemetrexed corresponds to a weighted average of 'stable' and 'responding' patients with no AEs, whereas the value applied to the placebo patients is that of 'stable with fatigue'. These choices are difficult to justify. The ERG prefers instead to use the full range of Nafees[15] utility values (including those with AEs) to derive a weighted average separately for each arm of the trial based on the frequency of condition (response/stable) combined with the AEs recorded in the trial. This approach yields a value of 0.6568 for pemetrexed patients compared to 0.6628 for placebo patients indicating the slight advantage anticipated for the placebo arm. Substituting these values into the model increases the submitted ICER from £33,732 to £36,798 per QALY.

Other utility values are drawn either from Nafees[15] or from a paper by Berthelot.[16] These are probably reasonable estimates. The last three weeks of life (the terminal cycle value) is treated separately and given a value of 0.47, which the ERG considers is probably too high. A value of about 0.2 might be more appropriate, but since it applies only to such a short period of time its influence on the size of the ICER is insignificant.

Chemotherapy costs

In the submitted model, CTX costs are estimated on the basis of the mean number of cycles of treatment per patient multiplied by an estimated average cost per cycle. In line with the ERG's view that the base case analysis should relate to the whole trial period, the trial mean of 7.978 cycles of pemetrexed in place of the manufacturer's truncated estimate of 5.84 cycles is used.

The mean cost per cycle of CTX in the model is based on the distribution of body surface area (BSA) of patients in the JMEN trial.[2] However, the trial involved 35% Asian subjects and did not include any UK patients. The ERG therefore considered it more appropriate to use a UK source for the distribution of BSA, leading to a slightly reduced acquisition cost. Further adjustments were required to include the cost of supplementation with dexamethasone, vitamin B12 and folic acid (omitted in the model), and to account for discounting of CTX cycles delivered beyond the first year of the trial period (i.e., beyond the 17 cycle maximum limit). In addition some minor adjustments were also made to the acquisition costs of second-line CTX (docetaxel and erlotinib) for compatibility with other lung cancer STAs. The net effect of these changes to the parameter values and discounting logic of the model is to increase the estimated ICER from £33,732 to £43,183 per QALY gained.

5.5.3 Additional issues noted by ERG

Continuity correction

A continuity correction is applied to models where a quantity is estimated at fixed time points, but the entity of interest (e.g. cost or survival) is accrued over the period between the fixed points. Relying only on values at either of the fixed points defining an interval may lead to systematic over or under-estimation. A traditional approach is to apply a 'half-cycle correction' during the first period to compensate for the inherent bias. However, this method is flawed as it is accurate only when the sequence of time periods extends indefinitely (e.g. to death), and if the quantities estimated are not subject to discounting. A more reliable method is the 'mid-cycle correction' where the entity of interest in each period is estimated from the average of the initial and final values for the period multiplied by the duration of the interval.

In the manufacturer's model, this issue does not apply to the estimation of costs which are not estimated through a Markov process. However, it is important for patient outcomes, and a 'half-cycle correction' has been applied to all survival estimates, including the base case estimate which combines the area under the Kaplan-Meier survival function during the trial period with a parametric estimate (exponential or Weibull) thereafter out to the six year time horizon. This method is not appropriate since the value of the survival obtained from the trial period is not subject to the same end-point bias as modelled estimates and should not be corrected in the first period. The correct approach is to use the area under the curve (AUC) from the trial analysis unaltered, and then calculate 'mid-cycle' corrected estimates for the remainder of the model duration derived from a parametric model. When this approach is applied to the manufacturer's base case, the incremental utility gain is reduced by 3.5%, and the ICER correspondingly increased to £34,860 per QALY gained.

Discounting

Pemetrexed CTX cycles are limited in the manufacturer's model to a maximum of 17 cycles per patient (i.e. less than 1 year), so no discounting is applied to maintenance CTX costs. If more than 17 cycles are considered (as in the full trial analysis (see 5.5.2) preferred as the ERG's base case) discounting must be used for cycles of pemetrexed beyond 17 cycles. Discounting has been included in the drug costs amendment described above in section 5.5.2.

Post progression costs and utilities are calculated in the submitted model by apportioning the overall mean survival between maintenance, second-line CTX, BSC and terminal care phases. Since apportioning is carried out on the basis of discounted overall survival estimates, and the costs are then discounted again, the post progression costs are double discounted in the model. In addition, the

estimation of QALYs relies on the same discounted survival values and were similarly double discounted, so that incremental utilities are also affected. The net consequence of correcting this error in the manufacturer's base case is to increase incremental costs by a small amount (for BSC costs only), but to increase incremental QALYs by about 5.5%, so that the ICER falls to £32,091 per QALY gained.

The discounting applied in the model to the four care phases is based on simplistic assumptions. All maintenance CTX cycles are assumed to occur in the first year (consistent with the imposed maximum cycles limit but not with the trial data), all terminal care is assigned to year three, all second-line CTX is placed in the first year, and all BSC is assumed to occur only in years one or two. These calculations are based on manipulation of mean numbers of cycles and take no account of the very skewed distributions of overall survival and progression free survival reported in the trial. Rectifying these inappropriate approximations, based on the pattern of treatment of patient experience in the trial, results in small reductions in both incremental costs and QALYs, with the manufacturer's base case ICER falling to £33,640 per QALY gained.

Monitoring costs

Patients on pemetrexed CTX receive monitoring/assessment contacts more frequently than 'watch and wait' patients involving out-patient visits and CT scans. The cost per additional CT scan is estimated as £112.54 (weighted average across all types of CT scan in NHS Reference Costs 2007/8, codes TCPDIAGIM_OP/RA08Z-RA14Z & RA50Z) and the cost per out-patient follow-up visit as £124 (Non-admitted multi-professional consultant led visit in NHS Reference Costs 2007/8, code TPCTLFUMFF800). Furthermore, UK practice does not involve scans and clinical assessment every 2 cycles (6 weeks) as used in the JMEN trial,[2] for either continuing CTX or for 'watch and wait' monitoring. To illustrate the differential effect of monitoring costs we assume that during 'watch and wait' following first-line CTX patients are scanned and examined after 3 months, 6 months, 12 months and every 6 months thereafter until progression. Similarly during active CTX it is assumed that patients are scanned and examined every 4 cycles (12 weeks) until progression. Including these additional costs in the submitted model increases the incremental cost per patient in the trial by £249, and increases the ICER to £34,651 per QALY gained.

Arithmetic error

A minor error has been detected in calculating the proportion of patients assumed to receive docetaxel and erlotinib in second-line therapy. When this is corrected the ICER for the manufacturer's base case rises slightly to £33,817 per QALY gained.

6 SUMMARY OF ADDITIONAL WORK BY ERG

Following an initial assessment of the manufacturer's submission documents and the submitted decision model, the ERG was concerned about two important issues:

i) When technology is assessed within the context of a sequence of treatments there is a potential risk of progressive casemix changes influencing the overall outcomes seen in a clinical trial. Most statistical techniques require broad assumptions of homogeneity within the studied population, and these can be violated when the intermediate events occurring to patients prior to the final outcome predispose some patients to better or worse prognosis than others in the trial. Of particularly importance in this instance is the assumption made by the model designers that overall survival is the same for patients receiving second-line CTX and those who do not, without consideration of the likelihood that patients not offered further treatment may be deemed of poorer health status and with worse prognosis.

ii) The manufacturer's submission indicated that the use of pemetrexed as maintenance therapy for NSCLC might be appropriate for consideration under the 'end of life' criteria for assessment. This requires consideration of the expected gain in life expectancy (which should be more than 3 months), and also of the extent of uncertainty associated with the effectiveness and cost-effectiveness results presented.

In view of these concerns, the ERG felt it was important to carry out more detailed analysis of the JMEN clinical trial results, in order to clarify the robustness of the patient outcome benefits claimed for pemetrexed. Further, since the manufacturer failed to present a probabilistic sensitivity analysis as recommended in the NICE Methods Guide,[23] the ERG considered that it may be important to the Appraisal Committee to have access to such an analysis.

For these reasons the ERG formally requested an anonymised data extract of individual patient data (IPD) records from the JMEN trial,[2] focussed on information relating to patient characteristics, key events and their timings to allow these issues to be addressed. This information was received by the ERG on 11 September 2009, together with other information requested in the letter of clarification (see Appendix 4 for details of clarification information received).

The ERG had already identified shortcomings in the manufacturer's approach to projecting patient survival beyond the trial period, and designed an alternative method more likely to yield reliable results (see section 5.5.2). In response to queries raised by the ERG about protocol violations, the manufacturer commented that:

"…two uncensored patients were prospectively identified in the regulatory submissions to the EMEA and FDA as "significant protocol violations" for having continued therapy beyond their progression dates, despite documented radiographic assessments every 2 cycles as mandated in the protocol."

Therefore these two cases were excluded from subsequent analyses.

The mean overall survival was estimated and compared for the two JMEN trial arms, using the approach described in section 5.5.2. For pemetrexed, the estimated mean OS is 22.21 months (21.59 months discounted), and for placebo 16.62 months (16.27 months discounted). The difference in favour of pemetrexed is then 5.58 months (0.465 years), or 5.46 months (0.455 years) discounted. The estimated 95% confidence interval for the undiscounted difference is 3.03 – 8.13 months.

The ERG also analysed the IPD in four subgroups defined by two criteria:

- trial medication (Pemetrexed + BSC vs. Placebo + BSC)
- use of second-line CTX (at least one further CTX agent used vs. no more CTX).

For each group the same survival analysis method was applied yielding the mean survival time recorded during the trial period (area under the curve from Kaplan-Meier analysis) and the mean projected survival from the end of the trial period to the time horizon on the decision model (six years). When combined to yield overall estimates of OS in each arm, the results were similar to those described above, though with much wider confidence limits. This indicates that there is no strong evidence of a 'survivor bias' effect distorting projected benefits. However, there is clear evidence of important differences in both pemetrexed and placebo arms of the trial between the estimated OS for patients receiving 2nd line CTX, and those who do not (contrary to the assumption of equivalence made in the manufacturer's model).

6.1 ERG base case results

New model results were generated by the ERG to take account of each of the issues previously identified (sections 5.5.2 and 5.5.3 above); the separate effect of each change is shown in the upper section of Table 6-1 compared to the manufacturer's submitted base case analysis. The most influential amendments are the removal of a limit on the number of cycles of treatment any patient could receive, and substitution of utility values based on the incidence of AEs reported in the JMEN trial.[2] The combined effect of these changes is to increase the incremental cost attributable to use of pemetrexed by 35% as well as reducing the incremental QALYs gained by 2%, so that the ICER increases from £33,732 to £47,239 per QALY gained.

In the lower part of Table 6-1 the results of re-analysing the trial IPD (revised OS estimates and recalculated mean number of cycles of maintenance treatment) are combined with the earlier changes. This has little effect on the estimated costs and increases the attributable QALYs in both arms, but differentially in favour of placebo, resulting in a further deterioration in the estimated ICER to £51,192 per QALY gained.

6.2 ERG's consideration of decision uncertainty

Since the manufacturer's base case yielded an ICER close to the upper limit of the range conventionally considered to be cost effective, the ERG originally considered attempting to rectify the absence of a PSA in the submitted model, and this was one of the reasons for requesting an extract of the trial IPD.

Sufficient time was not available to carry out a full PSA which involves specifying uncertainty parameters and distributions for all model variables. However, it proved possible to produce an approximate probabilistic analysis, based on two statistics obtained from analysis of the trial IPD:

- the estimated mean OS per patient

- the mean cycles per patient of treatment administered in each arm of the trial

The effect of variations in the number of treatment cycles on model estimates of the net cost per patient were described using a linear regression equation. Similarly, model estimates of total QALYs per patient were based on a multiple regression involving variations in both the number of treatment cycles and the mean OS. These relationships were then applied together with the relevant standard errors of the parameters to yield 1000 randomly generated probabilistic scenarios.

In Figure 6-1 the scatterplot on the cost-effectiveness plane is displayed, which indicates that the uncertainty in incremental benefit is more influential than that for incremental cost. The scatterplot indicates that in all scenarios results fall within the 'upper-right' quadrant signifying both increased cost and increased benefit from use of pemetrexed as maintenance therapy. The incremental cost is dominated by the additional cost of pemetrexed and its administration, and shows very limited variation.

Figure 6-2 shows the corresponding cost-effectiveness acceptability curve, with no measurable probability of cost effectiveness for a 'willingness to pay' threshold of £30,000 per QALY gained,

and 50% probability of cost effectiveness achieved for a threshold above about £51,000 per QALY gained.

As a result of the manufacturer's univariate SA, it was concluded that the key drivers of uncertainty in their analysis were:

- the number of cycles of CTX administered, and

- the utility values applied to health states.

In the light of the ERG's investigations, an additional issue is discussed:

- the method used to estimate survival gain, and

The MS considers the case for restricting the number of cycles of maintenance CTX to no more than 17 or even less (possibly 10). Although this would clearly reduce the extra cost of pemetrexed treatment, it is claimed that this would not reduce the benefits to patients – a claim which seems to lack any substantial basis.

The ERG previously commented on the extent of 'crossover' present in the JMEN trial[2] in that a greater proportion of placebo patients received second-line CTX than those randomised to pemetrexed. Moreover, it is clear that this excess is wholly accounted for by use of pemetrexed CTX as second-line therapy. It is argued by the manufacturer that this constitutes a bias which operates against the comparison of pemetrexed as maintenance and 'watch and wait'. However, adjustment to the model to reduce the number of patients in both trial arms by assuming that those who were given pemetrexed as second line CTX are instead given no second line CTX at all suggests that the ICER may move slightly against pemetrexed. It may be cautiously concluded that any potential crossover bias against pemetrexed is uncertain and likely to be of limited magnitude.

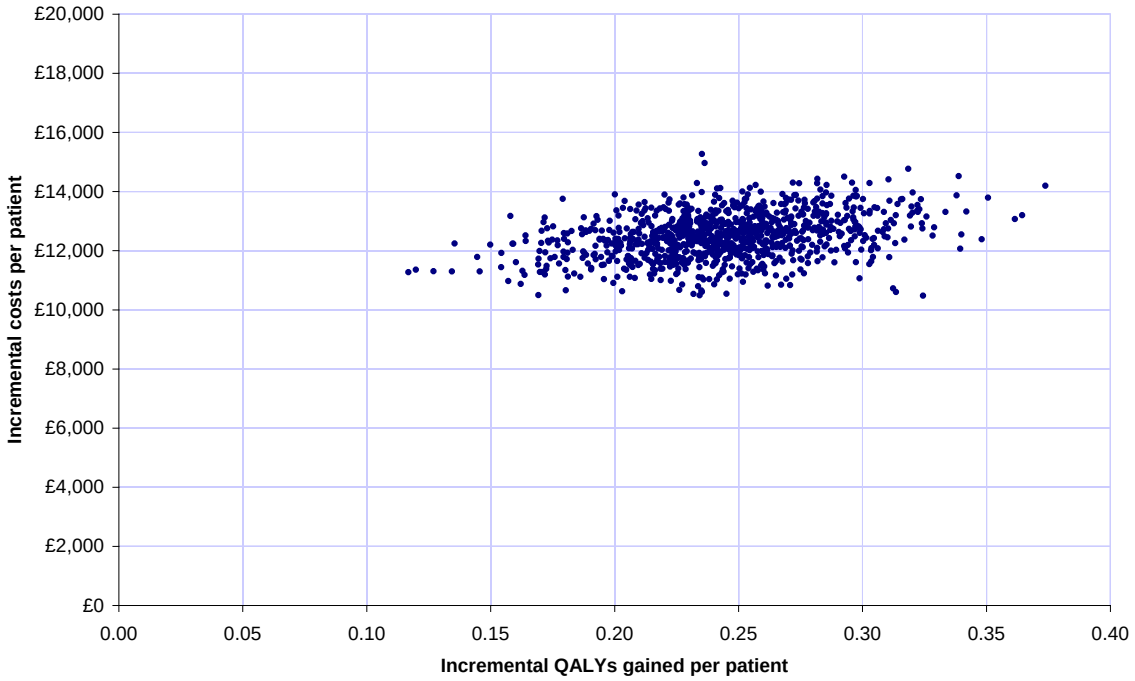

Figure 6-1 Cost-effectiveness scatterplot from approximate PSA for ERG preferred scenario using model amendments and revised survival analysis

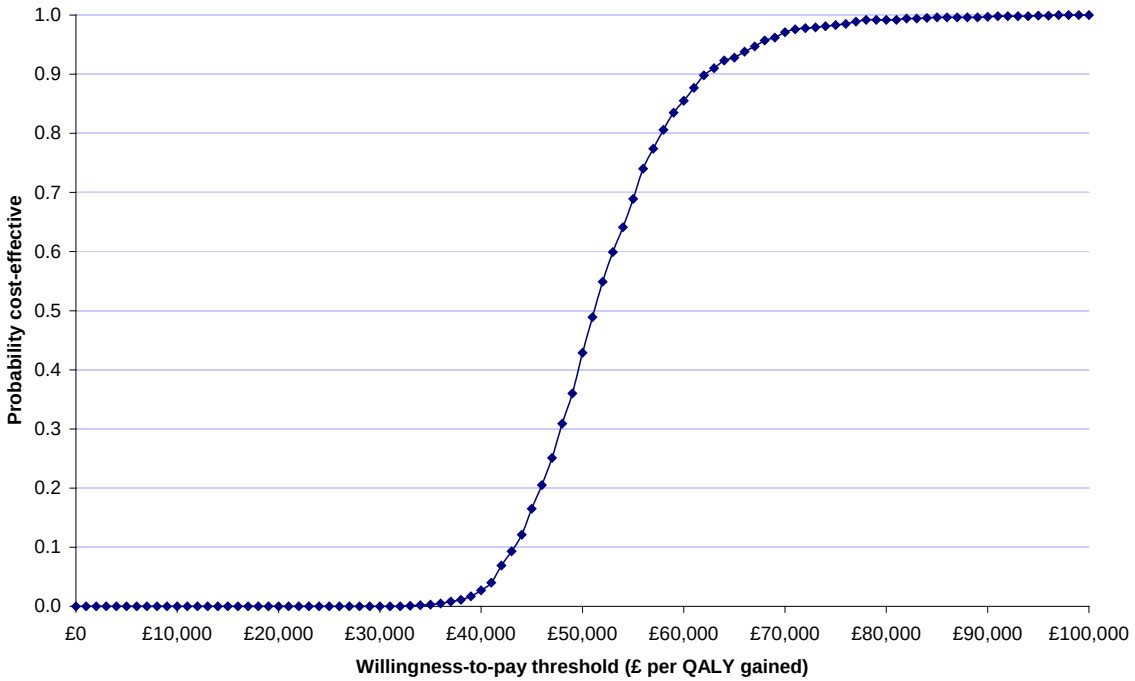

Figure 6-2 Cost-effectiveness acceptability curve scatterplot from approximate PSA for ERG preferred scenario using model amendments and revised survival analysis

Table 6-1 Effect of corrections and amendments made by ERG to the manufacturer's model for the non-squamous population

	Pemetrexed		Placebo		Incremental		ICER	Changes		
Model amendment	Costs	QALYs	Costs	QALYs	Costs	QALYs	(£/QALY)	Costs	QALYs	ICER
Submitted base case	**£17,455**	**0.9697**	**£8,318**	**0.6988**	**£9,137**	**0.2709**	**£33,732**	**-**	**-**	**-**
All cycles of pemetrexed and revised CTX costs	£20,638	0.9841	£8,323	0.6989	£12,315	0.2852	£43,179	+£3,178	+0.0143	+£9,447
Revised utility values	£17,455	0.9540	£8,318	0.7057	£9,137	0.2483	£36,798	-	-0.0226	+£3,066
Continuity correction	£17,405	0.9467	£8,288	0.6851	£9,117	0.2615	£34,860	-£20	-0.0094	+£1,128
Correct double discounting	£17,522	1.0006	£8,352	0.7149	£9,169	0.2857	£32,091	+£32	+0.0148	-£1,641
Discounting assumptions	£17,421	0.9617	£8312	0.6909	£9,109	0.2708	£33,640	-£60	-0.0001	-£88
Include monitoring costs	£17,838	0.9697	£8,452	0.6988	£9,386	0.2709	£34,651	+£249	-	+£919
Correct arithmetic	£17,398	0.9658	£8,248	0.6953	£9,149	0.2706	£33,817	+£12	-0.0003	+£85
Combined effect of above changes	**£20,925**	**0.9539**	**£8,370**	**0.6881**	**£12,555**	**0.2658**	**£47,239**	**+£3,418**	**-0.0051**	**+£13,507**
Combined effect of all changes including IPD survival analysis (excluding significant protocol violations)	**£20,902**	**0.9851**	**£8,382**	**0.7405**	**£12,520**	**0.2446**	**£51,192**	**+£3,383**	**-0.0263**	**+£17,460**

ICER= incremental cost effectiveness ratio; QALY=quality adjusted life year; CTX=chemotherapy; IP= individual patient data

6.3 Summary of economic evidence

6.3.1 Economic evaluation results

Base case: Manufacturer

- The manufacturer reports an ICER of £33,732 per QALY gained for the comparison of pemetrexed + BSC versus 'watch and wait' + BSC in the non-squamous population (exponential survival function).

- The manufacturer reports an ICER of £39, 364 per QALY gained for the comparison of pemetrexed + BSC versus 'watch and wait' + BSC in the adenocarcinoma subgroup (exponential survival function).

- Results of the SA conducted by the manufacturer suggest that, based on the assumptions and parameters explored, pemetrexed + BSC is likely to be cost-effective compared with 'watch and wait' + BSC.

Base case: ERG

- The ERG base case ICER was estimated at £51,192 per QALY. This figure is based on correcting a number of methodological errors (which together increase the ICER to £47,239 per QALY) and re-analysis of survival estimates excluding any cases of significant protocol violation from the JMEN trial[2] IPD.

- The ERG also undertook a basic PSA (to account for the fact that the manufacturer did not provide PSA results), and estimated that with a threshold of £30,000 per QALY, there is no measurable probability that pemetrexed is cost effective. If the threshold is increased to £51,000 per QALY there is about 50% probability of cost effectiveness being achieved.

6.3.2 Economic issues and uncertainties

- The costing of CTX treatment and the utility values used in the manufacturer's model are not ideal and underestimate the size of the ICER.

- The manufacturer's SA is limited and may not reflect the true uncertainty surrounding the decision problem.

- The cycle capping rule implemented in the manufacturer's model only affects costs and does not take account of any reduction in outcomes caused by capping the maximum number of cycles at 17 rather than allowing the JMEN trial[2] maximum of 55. Again, this capping rule underestimates the size of the ICER.

- Analysis of the JMEN trial[2] IPD and projection of survival from the end of the trial period by the ERG suggested a slight increase in the estimated mean OS benefit from 5.5 months to 5.58 months.

7 END OF LIFE CRITERIA

7.1 Introduction

This section provides an overview and critique of the manufacturer's case for pemetrexed as an end of life maintenance treatment for patients with NSCLC.[25] The NICE end of life treatments criteria[25] has three key points:

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

7.2 Application of the end of life treatment criteria

Each of the end of life criteria is discussed below for the case of pemetrexed for the maintenance therapy for patients with NSCLC.

7.2.1 Patient life expectancy of less than 24 months

The manufacturer makes a case that the OS of untreated patients with NSCLC is in the region of 7.9 – 10.3 months, far less than 24 months, using LUCADA[7] data and JMEN placebo data.[2] The ERG is of the opinion that the mean life expectancy of patients with stage IIIb/IV NSCLC is likely to be less than 24 months.

7.2.2 Life extension of at least three months

The JMEN trial[2] showed a mean OS benefit of 5.3 months in the pemetrexed maintenance arm compared with the placebo arm for the licensed non-squamous population. However, the JMEN trial[2] did not include any UK centres and included a large proportion of Asian patients who are known to do better than other ethnicities suffering with NSCLC; hence it is not certain that the benefit could be replicated in a UK trial setting.

ERG re-analysis of JMEN trial data including an alternative method for projecting survival beyond the trial period yielded an estimated mean gain in OS per patient treated with

pemetrexed of 5.58 months (3.03 – 8.13), supporting a life extension of greater than three months.

7.2.3 Licensed for a small patient population

The manufacturer uses the LUCADA 2007 audit,[7] statistics from Cancer Research UK,[4] and figures quoted in NICE literature[26 27] alongside Lilly trial/market research data to estimate the patient population for pemetrexed for maintenance therapy and the whole licensed population for pemetrexed.

Patient population for pemetrexed for maintenance therapy

In Table 9 of the MS (page 49) the MS estimates this population to be 949 patients, whereas in Figure 16 of the MS (page 125) the figure ranges from 1121 – 2165 patients. This arithmetic difference is likely to arise from the different methods used to estimate the number of patients who respond to first-line therapy and the number of patients who do not receive first-line pemetrexed. Neither of the calculations and assumptions used to derive the estimates is fully explained and therefore it is difficult to determine the accuracy of the figures.

Patient population for the whole pemetrexed licensed population

In Appendix 6 of the MS (Appendices, page 41) the manufacturer estimates the total licensed population to be 3426 patients, which equates to the sum of patients with first-line NSCLC, patients with second-line NSCLC, patients with mesothelioma and patients with NSCLC requiring maintenance therapy. It is worth noting that this estimate is based on the lower estimate of 949 pemetrexed maintenance therapy patients. If the higher estimate of 2165 pemetrexed maintenance therapy patients is used then the whole licensed population becomes 4642.

The derivation of the manufacturer's estimate of the whole licensed patient population, as with the calculation of the maintenance therapy population, relies heavily on the amalgamation of several data sources, which may or may not be compatible. The figures become particularly uncertain once the estimates of the proportion of patients who receive first-line CTX are examined. The manufacturer has used the LUCADA 2007[7] figure of 23% (exact figure 23.17%) to estimate the proportion of patients with NSCLC in England and Wales with advanced non-squamous carcinomas who receive first-line CTX. In fact, the 23%

in the LUCADA audit [7] refers to the proportion of lung cancer patients in England and Wales (of any stage or histology, including small cell lung cancer) who receive CTX as their first-line treatment. Thus the 23% is not reflective of patients with advanced NSCLC with non-squamous histology.

In 2004 the Royal College of Physicians worked with NICE to develop the CG 24 Lung Cancer Guidance and it was estimated that approximately 50% of patients with advanced NSCLC were eligible for CTX.[27] If the estimate of the proportion of patients who receive first-line CTX were to increase to 50%, then the number of maintenance therapy pemetrexed patients would double to approximately 2000 – 4000 patients, and the number of licensed population patients would increase to approximately 6000 – 9000.

A further uncertainty when attempting to estimate the patient population is how the recent NICE approval of pemetrexed for first-line CTX{NICE, 2009 #44} will affect the use of pemetrexed in the UK, both as a first-line therapy and as a maintenance therapy. It is not clear that the full impact of this approval has been accounted for by the manufacturer when estimating the size of the patient population. Again, between sections of the MS the estimates of patient numbers who receive first-line pemetrexed and who are eligible to receive pemetrexed maintenance therapy do not always concur (compare Figure 16 with Table 9 in the MS), which is a cause for concern.

The ERG is of the opinion that more information on the patient numbers is required, with transparent, consistent calculations and justifications for each step in order to estimate the size of the patient population.

7.3 Manufacturer's results

The manufacturer estimates a base case ICER for the non-squamous population of £33,732 per QALY (exponential function). This equates to a QALY weighting of between 1.12 (£30,000 threshold) to 1.69 (£20,000 threshold) based on the original HRQoL estimates.

The manufacturer also presents a case using revised HRQoL estimates based on a value of 0.8, which is the maximum utility value achievable by a person aged 55-64 (see MS page118). Using these revised QALY weighting figures the ICER reduces to £25,957 per QALY, which falls within the acceptable range of NICE cost-effectiveness thresholds (£20,000-£30,000).

7.4 ERG's results

The ERG base case ICER for the non-squamous population was estimated as £51,192 per QALY (see section 6) which equates to a QALY weighting of 1.7 - 2.6 based on original HRQoL estimates.

8 DISCUSSION

8.1 Summary of clinical-effectiveness issues

The manufacturer presents a case for the use of pemetrexed as a maintenance treatment for patients with non-squamous NSCLC whose disease has not progressed following four cycles of induction CTX with a platinum doublet containing gemcitabine, docetaxel or paclitaxel. The systematic review carried out by the manufacturer identified a single relevant RCT (JMEN[2]) which compares the use of pemetrexed + BSC with placebo + BSC. The ERG is confident that all published trial reports were identified.

The JMEN trial[2] included 663 patients with NSCLC. The population that furnishes the evidence for the clinical and cost-effectiveness evidence is a subset of the overall population of this trial, 481 patients with non-squamous histology. Some serious flaws relating to the conduct of the trial were identified by the ERG; these cast doubt on the results as reported by the manufacturer.

Firstly, the manufacturer reported that the primary outcome of the trial had been changed from OS to PFS; the ERG considered that the rationale for the change was not clearly justified, neither was the timing of the decision for the change clearly specified.

Secondly, histology was not a factor of the randomisation process and yet the key evidence in the MS relied on the results of the non-squamous subgroup.

A further issue with the trial was the lack of HRQoL data, which appears to be primarily due to poor trial management. This is not the first time that the ERG has reviewed a MS that contained few or no HRQoL data despite collection of this data being planned in the trial protocol. Whilst a patient's HRQoL experiences in a clinical trial may be of less importance to clinical investigators and their sponsors than the collection of primary outcome data, HRQoL is of key importance to the patients who undergo these treatments and to the clinicians whose job it is to administer them.

In terms of the generalisability of the trial to UK clinical practice, several factors were identified which may mean that the results of the trial, even if proved internally valid, may not be reproducible in the UK. In particular:

- The trial did not include any centres based in the UK;

- One third of the patients in the trial were of Asian origin, a group of patients known to have a more favourable prognosis in NSCLC. These patients appeared to have greater survival times (both treatment and placebo) compared to other non-Asian groups in the JMEN trial;[2]

- The patients in the trial received a large proportion of second-line agents (53% in the pemetrexed arm) that are not currently available in the UK. This may have influenced the size of the OS estimates observed in the trial and may mean the results would not be expected in UK practice;

- The trial excluded patients who received pemetrexed or vinorelbine as a first-line therapy, both of which are available in the UK;

- The trial allowed patients to receive unlimited cycles of maintenance therapy, which may not occur in clinical practice. It is uncertain if the same trial benefits would be obtained if the maximum number of cycles had been capped.

8.2 Summary of cost-effectiveness issues

The manufacturer's economic model also relied heavily on the JMEN trial[2] data together with an extrapolation component which took the 29 month trial data out to six years. The manufacturer's base case analysis appeared to indicate economic results close to the borderline of conventional cost-effectiveness acceptability (£33,732 per QALY for the non-squamous population). However, examination of the submitted model identified a number of errors and inconsistencies which, once corrected, increased the size of the ICER to £47,239 per QALY.

During the STA process the ERG requested access to an extract from the clinical trial IPD, which has proved invaluable in allowing a number of key issues to be examined in a way which would not otherwise have been possible. As a consequence, new estimates of OS gain (slightly greater than in the manufacturer's base case) were calculated which, together with the other model changes identified, led to a further important increase in the size of the ICER for use of pemetrexed in maintenance therapy to a level not normally considered cost effective (£51,192 per QALY).

The ERG also undertook a basic PSA (to account for the fact that the manufacturer did not provide PSA results), and estimated that with a threshold of £30,000, there is no measurable

probability that pemetrexed is cost effective. If the threshold is increased to £51,000, there is about 50% probability of cost effectiveness being achieved.

8.3 Implications for research

Due to the flaws within the JMEN trial[2] and the many factors which limit its generalisability to UK clinical practice, there is a need for further RCTs of pemetrexed for the maintenance treatment of patients with NSCLC. Such a trial should be limited to non-squamous patients only, and be randomised according to histology group, ideally with pre-planned subgroups of patients with large cell and adenocarcinoma. Preferably a trial should contain a substantial proportion of UK centres, which include patients reflective of UK clinical practice in terms of their age, performance status, and induction therapy. Any trial should also aim to capture HRQoL data much more effectively than is currently the norm in trials of advanced cancer. However, undertaking such a trial would be very costly and time consuming, and may not represent the best use of NHS resources.

There is also a need for more information on the role of maintenance therapy in NSCLC in general. Several important trials of CTX agents for the maintenance therapy of NSCLC are underway, notably SATURN[28] (which compares erlotinib with placebo) and ATLAS[29] (which compares erlotinib with bevacizumab and erlotinib). Neither of these trials is published at the time this ERG report was written, but the trials are due to report shortly. A further trial of pemetrexed for maintenance therapy following induction with pemetrexed is also underway (see MS pg 4), but results from this trial are not due until 2012. Together these trials may help to elucidate further the role of maintenance therapy for patients with NSCLC.

Given the fact that maintenance therapy is a new addition to the treatment pathway of patients with NSCLC in the UK, and that several RCTs offering discrete comparisons of various CTX options are due to report shortly, research within the UK may be best focussed on collating all of this information and undertaking a systematic review and economic evaluation of all maintenance therapy options for NSCLC.

9 REFERENCES

1. EMEA. Committee for medicinal products for human use post authorisation summary of positive opinion for Alimta. 2009; Available from: http://www.emea.europa.eu
2. Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomised, double-blind, phase 3 study. Lancet. 2009; Epub ahead of print [Sept 18].
3. Ou S-HIMDP, Ziogas AP, Zell JADOMPH. Asian Ethnicity Is a Favorable Prognostic Factor for Overall Survival in Non-small Cell Lung Cancer (NSCLC) and Is Independent of Smoking Status. Journal of Thoracic Oncology. 2009; 4(9):1083-93.
4. Cancer Research UK. Cancer stats. 2009 [07/2009]; Available from: http://www.cancerresearchuk.org/.
5. National lung cancer audit. Key findings about the quality of care for people with lung cancer in England and Wales incorporating headline and completeness data from Scotland. Report for the audit period 2006.
6. Gabrielson E. Worldwide trends in lung cancer pathology. Respirology. 2006; 11:533-8.
7. National lung cancer audit. Key findings about the quality of care for people with lung cancer in England and Wales incorporating headline and completeness data from Scotland. Report for the audit period 2007.
8. Clegg A, Scott D, Sidhu M, Hewitson P, Waugh N. A rapid and systematic review of the clinical and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small cell lung cancer: a systematic review. Health Technology Assessment. 2001; 5(32).
9. NICE. Lung cancer (non-small cell) - pemetrexed first line, single technology appraisal. 2009 [08/2009]; Available from: http://guidance.nice.org.uk/TA/Wave18/25.
10. NICE. The diagnosis and treatment of lung cancer. 2005; Available from: http://www.nice.org.uk/nicemedia/pdf/cg024fullguideline.pdf.
11. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in previously treated non-small-cell lung cancer. New England Journal of Medicine. 2005; 353(2):123-32.
12. Eli Lilly. Date on File, Market Research Data. 2008.
13. Lung cancer symptom scale. [cited 2009 August]; Available from: http://www.lcss-ql.com/.
14. Committee for medicinal products for human use (CHMP). Reflection paper on methodolical issues in confirmatory clinical trials with flexible design and analysis plan. London, March 2006; Available from: http://www.emea.europa.eu/pdfs/human/ewp/245902en.pdf.
15. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008; 6(84).
16. Berthelot JM, Will BP, Evans WK, Coyle D, Earle CC, Bordeleau L. Decision framework for chemotherapeutic interventions for metastatic non-small-cell lung cancer. Journal of the National Cancer Institute. 2000; 92(16):1321-9.
17. MIMS. 2009; Available from: http://www.mims.co.uk/.
18. Department of Health. NHS reference costs 2007-2008. 2008; Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_098945.

19. NICE. Guidance on Cancer Services: Improving Supportive and Palliative Care for Adults with Cancer. 2004; Available from: http://guidance.nice.org.uk/CSGSP/Guidance/pdf/English.
20. NICE. Lung cancer (non-small cell) - erlotinib, single technology appraisal. 2006 [07/2009]; Available from: http://guidance.nice.org.uk/TA162.
21. Duran A, Eggington S, Aristides M, Samyshkin Y. Unit cost and patterns of treatment of adverse events and providing best supportive care in patients with Stage IIB/IV non-small cell lung cancer. A report from opinion-based survey for the UK. : IMS Health 2008.
22. Hanna N, Shepherd FA, Fossella FV, Pereira JR, Demarinis F, Von Pawel J, et al. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. Journal of Clinical Oncology. 2004; 22(9):1589-97.
23. NICE. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2008; Available from: http://www.nice.org.uk/pdf/TAP_Methods.pdf
24. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working. BMJ. 1996; 313(7052):275-83.
25. NICE. Supplementary Advice to the Appraisal Committees - End of Life Treatments. . Available from: http://www.nice.org.uk/aboutnice/howwework/devnicetech/endoflifetreatments.jsp?domedia=1&mid=88ACDAE5-19B9-E0B5-D422589714A8EC6D.
26. NICE. The diagnosis and treatment of lung cancer: Clinical guideline 24. 2005 [02.2005]; Available from: http://www.nice.org.uk/nicemedia/pdf/cg024fullguideline.pdf.
27. NICE. Lung cancer (non-small-cell) - pemetrexed (maintenance): final scope. 2009; Available from: http://guidance.nice.org.uk/TA/WaveCRS2/47/FinalScope.
28. F Cappuzzo, T Ciuleanu, L. Stelmakh, S. Cicenas, A. Szczesna, E. Juhasz, et al. SATURN: A double-blind, randomized, phase III study of maintenance erlotinib versus placebo following nonprogression with first-line platinum-based chemotherapy in patients with advanced NSCLC. J Clin Oncol: 27:15s, 2009 (suppl; abstr 8001) ASCO; 2009.
29. V. A. Miller, P. O'Connor, C. Soh, F. Kabbinavar, for the ATLAS Investigators. A randomized, double-blind, placebo-controlled, phase IIIb trial (ATLAS) comparing bevacizumab (B) therapy with or without erlotinib (E) after completion of chemotherapy with B for first-line treatment of locally advanced, recurrent, or metastatic non-small cell lung cancer (NSCLC). J Clin Oncol: 27:18s, 2009 (suppl; abstr LBA8002) ASCO; 2009.

10 APPENDICES

Appendix 1: Validity assessment of the JMEN trial

Table 10-1 The manufacturer's approach to validity assessment and the ERG critique

NICE evaluative criteria	JMEN trial[2]	ERG comments
How was allocation concealed?	Allocation concealment was ensured as randomisation for all sites involved in the study was undertaken using a computerised, interactive, voice-activated response system at a central location. An unblinded pharmacist obtained the patient's treatment assignment from this system; investigators were thus shielded from knowledge of treatment assignment	The trial was conducted to a high standard in terms of allocation concealment and blinding
What randomisation technique was used?	Patients were randomised to pemetrexed or placebo in a 2:1 ratio and a minimisation principle was adopted to balance patient assignment between study arms	This is appropriate methodology. However, the evidence provided in the MS is pertinent to the subgroup of patients with non-squamous histology. Histology was not factored into the randomisation process, therefore the results are not derived from a randomised group
Was a justification of the sample size provided?	A sample size of approximately 660 patients was initially selected to provide analysis of OS with 80% power using a one-sided α level of 0.025, assuming 475 events and an OS HR of 0.767. The primary endpoint of the trial was later changed to PFS but nearly identical statistical assumptions and error control were maintained	The manufacturer clearly stated the sample size calculation for the whole trial population, but not for the non-squamous subgroup which provides the key evidence in the submission. The ERG considers that the change in primary endpoint was not fully justified and that the timing of the decision to change was unclear
Was follow-up adequate?	Each patient underwent a treatment period and a follow-up period. The treatment period consisted of treatment cycles, each 21 days long, administered until disease progression. The follow-up period began when the patient discontinued study treatment; follow-up included periodic tumour response evaluation until objective disease progression. Investigators followed all patients until death or study closure	Follow-up appears to be adequate. Data for QoL poorly followed-up
Were the individuals undertaking the outcomes assessment	Patients in the pemetrexed-treated arm were given pemetrexed 500 mg/m^2 via intravenous infusion on day 1 of a 21-day	The blinding processes in the trial were robust. As noted in the MS, it is unusual to have

aware of allocation?	cycle. Patients in the placebo-treated arm received an intravenous infusion of normal saline; to maintain blinding the pemetrexed and saline infusions were prepared by an unblinded pharmacist/designee at each site such that the preparations were visually indistinguishable. Unblinding was permitted if, in the opinion of the investigator, knowledge of treatment assignment would alter the management of a serious adverse event, otherwise physicians and patients were unblinded only at the time of disease progression	blinding in an oncology study; In addition, independent monitoring of investigators' outcome assessments was undertaken and reported
Was the design parallel-group or crossover? Indicate for each crossover trial whether a carry-over effect is likely	JMEN was a parallel-group study. However, patients who had disease progression were unblinded to study treatment and subsequent treatment was permitted at the discretion of the investigator, so some crossover did occur. Fewer patients in the pemetrexed arm received post-discontinuation therapy compared to placebo (53.2.5% vs. 67.3%, p<0.001). The rate of crossover from placebo to pemetrexed was 18.5%. Survival results are not likely to have been influenced by post-study therapy given the higher rate of follow-up treatment on the placebo arm, low rate of crossover, and the balanced selection of therapies between arms	It appears from the clarification provided by the manufacturer that the 'crossover' referred to is second-line treatment for the patients in the placebo arm. If this is the case, then removing the 18% of patients who 'crossed over' from the calculation of patients continuing to post-study therapy means that equal numbers of patients in both groups received post-study therapy. The IPD analysis indicates that this did not affect OS estimates
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 UK practice?	The JMEN trial[2] was a parallel group trial conducted at 83 investigational sites in 20 countries (Australia, Austria, Brazil, Bulgaria, China, Croatia, the Czech Republic, Germany, Greece, Hungary, India, Italy, Korea, the Netherlands, Poland, Romania, Spain, Taiwan, Turkey and the United States). There were no centres in the UK. However, the study design ensures that the trial results are very much relevant to the UK. The trial population is representative of patients with NSCLC as a whole since the inclusion/exclusion criteria for the JMEN trial[2] was such that only patients with locally advanced or metastatic NSCLC were enrolled. The patients received induction regimens similar to what the average NSCLC patient would receive in the UK, ie cisplatin or carboplatin in combination with gemcitabine, docetaxel and paclitaxel. The comparator in the JMEN trial[2] is placebo (watch and wait) plus	It is unfortunate that no centres in the UK were included in the trial. Subgroup analyses (post hoc) provided by the manufacturer suggest that the relative difference between trials between regions was similar to the total non-squamous population; however, OS for patients from the Asian region appears to be substantially greater than EU and non-Asian patients in both treatment and placebo arms

	BSC, which is the standard of care in the NHS	
How do the patients included in the RCT 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.	Patients in JMEN[2] were generally younger compared to the average NSCLC patient in the UK (LUCADA 2007). This was due to the inclusion criteria for the trial, which restricted patient entry to limit confounding factors. However, PS rather than age is a prognostic factor for OS in NSCLC, and so this is unlikely to impact the relevance of JMEN[2] results to UK patients. More patients in JMEN have adenocarcinoma and fewer patients have NSCLC-NOS than seen in LUCADA. This is due to better diagnosis in clinical trial compared to usual care. The proportion of patients with adenocarcinoma in the UK is likely to increase with improvements in diagnostic specificity over time. Most patients in JMEN[2] were of good performance status (PS 0-1). In LUCADA, 34% of patients were of good performance status. As mentioned previously, patients in LUCADA include those with lung cancer in general, irrespective of lines of treatment or eligibility for chemotherapy and so these patients are not necessarily representative of the average patient who would receive pemetrexed maintenance treatment, since in actual clinical practice, only patients who are relatively fit would receive chemotherapy	The trial limits the patient group to those with good performance status. This is appropriate; however, these patients are in the minority in clinical practice in England and Wales. The trial protocol differed from UK practice in a number of different ways. These are listed below Patients were allowed to have unlimited cycles of maintenance. This will not happen in UK practice, hence it is impossible to know what OS benefit will occur with fewer cycles. The patients had fewer first-line CTX than UK practice A substantial proportion of patients in the trial were of Asian ethnicity; this group are known to have a better prognosis. This group appeared to do better in the JMEN trial.[2] Patients received a number of second-line treatments that are not used in the NHS; 53% of treatments in the treatment arm are not available along with 36% in the placebo arm. All patients in the trial were of good performance status – not most as stated here by the manufacturer
For pharmaceuticals, what dosage regimens were used in the RCT? Are they within those detailed in the Summary of Product Characteristics?	See Section 6.3.1 for dosage regimens. These were as per the SPC for pemetrexed	These were appropriate
Were the study groups comparable?	The study groups were well balanced in terms of prognostic factors and other baseline characteristics and histology	The study groups were comparable
Were the statistical	See Section 6.3.5 for a description of the	The analyses were appropriate

analyses used appropriate?	statistical analysis for JMEN	
Was an intention-to-treat analysis undertaken?	Yes. ITT was undertaken for efficacy and safety analysis	ITT was undertaken
Were there any confounding factors that may attenuate the interpretation of the results of the RCT(s)?	None known	As noted previously, randomisation in the JMEN trial[2] did not include histological status although the key evidence in the submission is derived from a histological subgroup It is not clear when and why the primary endpoint in the JMEN trial[2] was changed from OS to PFS. Patients in the JMEN trial[2] received either docetaxel, gemcitabine or paclitaxel as induction therapy. In England and Wales, patients may also receive vinorelbine or pemetrexed. Patients in the JMEN trial[2] received a number of post-study therapies not given to patients in England and Wales

Appendix 2: Manufacturer's analyses of JMEN trial results by region

Note that these are the results of analyses unadjusted for potential differences in other prognostic factors.

Table 10-2 Clinical effectiveness results of the JMEN trial[2] by region (non-squamous population)

	N	Unadjusted hazard ratio (95% CI)	Log-rank p-value	Pemetrexed	Placebo
EU region	230				
Overall survival		0.67 (0.49-0.92)	0.014		
Median (months)				13.8	8.1
One-year rate				55%	36%
Progression-free survival	231	0.48 (0.35-0.66)	<0.00001		
Median (months)				4.6	2.7
Non-Asian region	310				
Overall survival		0.67 (0.51-0.88)	0.004		
Median (months)				13.0	8.5
One-year rate				54%	36%
Overall survival from start of induction		0.68 (0.51-0.89)	0.006		
Median (months)				16.2	12.0
One-year rate				67%	49%
Asian region	171				
Overall survival		0.75 (0.51-1.10)	0.139		
Median (months)				18.9	13.8
One-year rate				71%	54%
Overall survival from start of induction		0.75 (0.51-1.10)	0.138		
Median (months)				21.9	17.1
One-year rate				81%	67%

EU region = Austria, Bulgaria, Croatia, Czech Republic, Germany, Greece, Hungary, Italy, Netherlands, Poland, Romania, Spain
Non-Asian region = EU region, Australia, Brazil, Turkey, US
Asian region = China, India, Korea, Taiwan

Appendix 3: Manufacturer's analyses of JMEN trial results by subgroup

Table 10-3 Covariate-adjusted final OS and PFS for patients with non-squamous histology in the JMEN trial[2]

Variable	**N (number of events)**	**HR (95% CI)[a]**	**p value[c]**
Overall Survival	474 [a,b] (328)		
Pemetrexed v placebo		0.70 (0.56-0.88)	0.0021
ECOG performance status (0 v 1)		1.23 (0.98-1.54)	0.0792
Cisplatin[d] (yes v no)		1.04 (0.81-1.33)	0.7816
Induction response (PR/CR v SD)		0.97 (0.77-1.22)	0.7721
East Asian (yes v no)		0.69 (0.53-0.90)	0.0066
Non-smoker (yes v no)		0.90 (0.67-1.20)	0.4741
Gender (male v female)		0.66 (0.50-0.87)	0.0035
Age (<65 years v >65 years)		0.87 (0.69-1.10)	0.2537
Stage (IIIB v IV)		1.12 (0.84-1.49)	0.4587
Progression Free Survival	475[a e] (359)		p value[f]
Pemetrexed v placebo		0.45 (0.36-0.56)	<0.0001
ECOG performance status (0 v 1)		1.04 (0.84-1.29)	0.725
Induction response (PR/CR v SD)		1.04 (0.83-1.30)	0.739
East Asian (yes v no)		1.12 (0.87-1.42)	0.383
Non-smoker (yes v no)		1.02 (0.78-1.34)	0.861
Gender (male v female)		0.77 (0.59-0.99)	0.040
Age (<65 years v >65 years)		1.19 (0.94-1.50)	0.153

[a] Stratified by non-platinum component of induction therapy (gemcitabine v paclitaxel/docetaxel)
[b] Seven patients excluded due to missing values for one or more co-factors
[c] p-value from chi-square test
[d] Description of platinum agent in induction regimen: all patients treated with a platinum-based regimen, either cisplatin (yes) or carboplatin (no)
[e] Nine patients excluded due to missing values for one or more co-factors
[f] p-value taken from Mantel Haenzel square test

Table 10-4 Summary of efficacy parameters by subgroups, pemetrexed v placebo for the patients with non-squamous histology in the JMEN trial

Subgroup	**N**	**Overall survival**		**Progression free survival**	
Stage		Unadjusted HR (95% CI)	Median (months)	Unadjusted HR (95% CI)	Median (months)
Stage IIIB	85	0.52 (0.31-0.87)	17.5 vs. 8.7	0.56 (0.33-0.94)	4.3 vs. 1.6
Stage IV	395	0.75 (0.58-0.97)	15.0 vs. 10.6	0.43 (0.33-0.55)	4.8 vs. 2.7
Response					
Partial response to induction therapy	221	0.83 (0.59-1.15)	14.4 vs. 11.7	0.45 (0.32-0.61)	4.6 vs. 1.7
Partial or complete response* to induction therapy	226	0.81 (0.58-1.12)	14.4 vs. 11.7	0.45 (0.33-0.61)	4.5 vs. 1.7
Stable disease with induction therapy	252	0.61 (0.45-0.83)	16.6 vs. 8.6	0.44 (0.32-0.61)	4.5 vs. 2.8
Induction therapy					
Gemcitabine/cisplatin induction therapy	168	0.84 (0.57-1.24)	13.8 vs. 11.0	0.48 (0.33-0.70)	4.2 vs. 2.8
Gemcitabine/carboplatin induction therapy	127	0.75 (0.48-1.17)	14.0 vs. 9.1	0.55 (0.36-0.84)	4.6 vs. 1.6
Paclitaxel/carboplatin induction therapy	125	0.60 (0.39-0.94)	16.5 vs. 9.1	0.41 (0.26-0.64)	4.7 vs. 2.8
Paclitaxel/platinum* induction therapy	157	0.65 (0.44-0.96)	16.5 vs. 10.3	0.43 (0.29-0.65)	4.6 vs. 2.8
Taxane*/platinum* induction therapy	185	0.57 (0.40-0.82)	16.6 vs. 9.1	0.36 (0.25-0.53)	4.8 vs. 2.6
Carboplatin/cisplatin					
Cisplatin induction therapy	208	0.80 (0.56-1.12)	14.0 vs. 11.5	0.48 (0.35-0.68)	4.1 vs. 2.8
Carboplatin induction therapy	272	0.62 (0.46-0.83)	15.9 vs. 8.8	0.42 (0.31-0.57)	5.0 vs. 2.3
Performance status					
Performance status 0	193	0.57 (0.39-0.82)	17.7 vs. 10.3	0.33 (0.23-0.48)	5.5 vs. 1.6
Performance status 1	286	0.80 (0.60-1.06)	14.1 vs. 10.6	0.53 (0.40-0.70)	4.3 vs. 2.8

HR = hazard ratio

*Combination of requested subgroups due to small sample sizes of individual groups

Appendix 4: Clarification items requested, responses received and ERG assessment

Clarification requested	Rationale	Manufacturer's response	ERG assessment
Subgroups Please provide overall survival (OS) and progression free survival (PFS) hazard ratios together with confidence intervals, and the actual OS and PFS figures for the licensed non-squamous population for each of the following subgroups by trial arm: • Disease stage (presenting outcomes for stage IIIb separately from stage IV) • Response status prior to maintenance therapy (presenting outcomes for patients assessed as complete response at the start of maintenance, separately from partial response patients and again separately for stable disease patients) • First-line treatment (presenting outcomes according to the first-line regimen – with gemcitabine/cisplatin, docetaxel/cisplatin, paclitaxel/cisplatin gemcitabine/carboplatin, docetaxel/carboplatin and paclitaxel/carboplatin patients analysed separately) • First-platinum treatment (cisplatin separately from carboplatin) • ECOG performance status (PS0 separately from PS1)	The exploration of subgroups in the MS was limited. However, the CSR indicates some possible differences between subgroups of patients, but only for the whole trial population (i.e. not the licensed population)	Full details of the requested analyses were presented	Significant subgroup differences only apparent for gender (PFS and OS) and East Asian ethnicity (OS only)
Second-line therapy a) Please provide a breakdown of second-line therapy (for the licensed non-squamous population) by trial arm, explaining the reasons for second-line therapy (whether progression or adverse events or other reasons). b) Please provide further clarification and justification of the 18.5% cross over reported in the submission (did cross-over always occur after unblinding, and did it always count as second-line treatment?).	The trial showed a difference between the rates of second-line chemotherapy – 67% in the placebo arm and 53% in the pemetrexed arm (which are used in the manufacturer's economic model). Looking more closely at these figures, 18.5% of the 67% of placebo patients receiving second-line chemotherapy received pemetrexed (compared to <1% in the pemetrexed arm). The CSR appears to indicate that placebo	a) Further information provided. Reasons for initiating second-line therapy were not collected, but a table of reasons for discontinuation of study medications was provided. b) Investigator had discretion under the trial protocol. c) Additional information provided. d) Duration of second-line therapy was not recorded in the trial	Crossover remains a potentially important complication in the interpretation of the trial results

Clarification requested	Rationale	Manufacturer's response	ERG assessment
c) Please provide a breakdown of second-line therapy by stage of disease for the licensed non-squamous population. d) Please also provide the mean and maximum number of second-line chemotherapy cycles for each trial arm for the licensed non-squamous population.	patients received the 18.5% pemetrexed second-line due to cross over. This is a trial specific occurrence and therefore not likely to be replicated in clinical practice		
Analysis by geographic region Please provide the results of any analyses undertaken by geographical region or centre for the licensed non-squamous population.	None of the centres in the trial were UK-based. The CSR indicated that analyses of results by region were to be undertaken	Additional effectiveness results provided	The data provided indicated that the patients from the Asian 'region' had greater OS than patients from non-Asian regions. This was true for the placebo and treatment arms. Relative OS matched that of other regions. It should be noted that the analyses were unadjusted for potential differences in other prognostic factors
Reasons for discontinuation Please provide information on reasons for discontinuation for the licensed non-squamous population for each trial arm	Issue raised by NICE	Additional information provided	N/A
Individual patient data To allow for a probabilistic sensitivity analysis to be undertaken, please provide a limited anonymised extract of the individual patient data from the JMEN trial[2] for each non-squamous patient as follows: • unique anonymised patient identifier • trial arm (pemetrexed or placebo) • days from randomisation to disease progression/withdrawal or censoring re-progression/withdrawal • censoring for progression/withdrawal (yes/no) • days from randomisation to death or censoring re-death • censoring for death (yes/no)	The submitted base case yields ICERs close to the upper range of cost-effectiveness, and one-way sensitivity analysis suggests that the result may be particularly sensitive to parameter uncertainty, especially as it relates to survival estimation and projection beyond the trial. The ERG notes that the manufacturer did not provide for a probabilistic sensitivity analysis in their economic model. It is likely that the members of the Appraisal Committee may wish to see results of such an exercise, or at least may seek the views of the ERG as to the likely findings from a PSA. Furthermore, analysis of patient populations	IPD extract provided electronically	Analysis of IPD proved especially helpful in carrying out additional survival analysis with assessment of associated uncertainty. Due to lack of time a full PSA could not be undertaken

Clarification requested	Rationale	Manufacturer's response	ERG assessment
• cycles of trial medication administered • cycles of second-line chemotherapy administered • type of second-line chemotherapy administered (list agent(s) or state "none") • days from randomisation to start of second-line chemotherapy • disease stage at baseline (IIIB/IV) • performance status at baseline (PS0/1) • histological sub-type (adeno/large cell/other) • response status prior to maintenance (complete response/partial response/stable disease)	subject to multi-stage therapies are particularly prone to case-mix distortion and bias when patients move between treatment stages		
Anti-emetic therapy Please provide the following for the licensed non-squamous population and for each trial arm: • medications prescribed • duration of treatment for each episode • number of patients given anti-emetic therapy at any time • total number of anti-emetic treatment episodes (or the total number of patient cycles in which treatment was given)	Since the incidence of nausea/vomiting differs between trial arms, information is required concerning the type and volume of treatments dispensed during the trial	Information was provided concerning the volume of use of anti-emetics during the trial, and the associated costs	Differences between the trial arms were small and unlikely to influence the results of the economic evaluation
Dose reduction Please provide the following for the licensed non-squamous population and for each trial arm: • total number of planned cycles of trial medication • total number of planned cycles where 100% of the planned dose was given • total number of planned cycles where 75% of the planned dose was given • total number of planned cycles where 50% of the planned dose was given • total number of planned cycles where none of the planned dose was given (i.e. missed cycles)	It is not clear from the manufacturer's submission how often dose reductions occurred in the trial	The requested information was provided	Dose reductions only occurred in about 1% of pemetrexed doses, and was considered too small to influence the results of the economic evaluation

Clarification requested	Rationale	Manufacturer's response	ERG assessment
Hospitalisations a) Please provide the summary information given in Table JMEN.12.10 of the Clinical Study Report for the licensed non-squamous population. b) Please provide further details of the hospitalisations the licensed non-squamous population as follows: • time/cycle in which episode occurred • length of stay • description and HRG/DRG code for the episode • any AEs related to the episode	It is not clear from the manufacturer's submission how often hospitalisations occurred in the trial	a) Some information was provided, but text stated that practice re admissions was very variable between countries and unlikely to be representative of UK practice b) Little information provided	Insufficient information to form a judgement concerning the frequency and cost of hospitalisations
Adverse events a) Please provide the number of episodes of toxicity as well as the number of patients suffering at least one episode (or the number of patient cycles involving an episode) for the licensed non-squamous population. b) Table 12 of the Manufacturer's submission (MS, page 52) references the file "DOF_JMEN_grade3/4AEs_ITT_non-squamous" but this file is missing from the documentation provided. Please provide this table.	Information provided in the MS and the CSR relate to patients only, and not to episodes (required for costing).	Requested information was provided	It appears that multiple episodes per patient were rare in the trial and so no amendments to the model are necessary
Transfusions Please provide information for the licensed non-squamous population together with the total number of each type of transfusion given (i.e. where a patient receives multiple transfusions).	Table JMEN.12.13 of the CSR shows the number of patients receiving transfusions for the whole trial population. This information is required for the licensed non-squamous population	A summary table was provided	Transfusions occurred rarely and usually only one per patient. The exception is packed red blood cells where an average of two transfusions were required per episode. No action required
Type of scan patients received in the trial Please provide information on the proportion of patients receiving chest-x ray, MRI and CT scan for the licensed non-squamous population and for each trial arm.	To consider the appropriate unit costs to apply to radiological investigations	A summary table was provided	The dominant mode of investigation was CT scan. No change required to the model