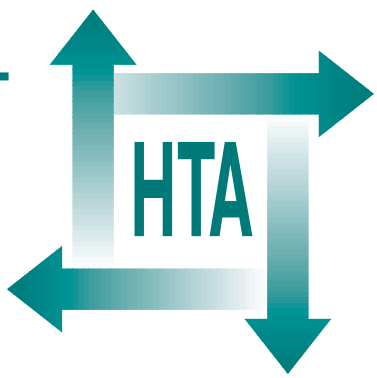
Health Technology Assessment 2001; Vol. 5: No. 32

Rapid review

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer

A Clegg
DA Scott
M Sidhu
P Hewitson
N Waugh

Health Technology Assessment
NHS R&D HTA Programme



How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer

A Clegg*
DA Scott
M Sidhu
P Hewitson
N Waugh

Southampton Health Technology Assessments Centre,
Wessex Institute for Health Research and Development,
University of Southampton, UK

* Corresponding author

Competing interests: none declared

Published May 2002

This report should be referenced as follows:

Clegg A, Scott DA, Sidhu M, Hewitson P, Waugh N. A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer. *Health Technol Assess* 2001;**5**(32).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA website (see opposite).

Part of the work described in this report has also been published as:
Clegg A, Scott DA, Hewitson P, Sidhu M, Waugh N. Clinical and cost effectiveness of paclitaxel, docetaxel, gemcitabine, and vinorelbine in non-small cell lung cancer: a systematic review. *Thorax* 2002;**57**:20–8.

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

The research reported in this monograph was commissioned by the HTA Programme on behalf of the National Institute for Clinical Excellence (NICE). Rapid reviews are completed in a limited time to inform the appraisal and guideline development processes managed by NICE. The review brings together evidence on key aspects of the use of the technology concerned. However, appraisals and guidelines produced by NICE are informed by a wide range of sources.

The research reported in this monograph was funded as project number 00/14/01.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme, NICE or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

R

Contents

Glossary and list of abbreviations

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader. In some cases, usage differs in the literature, but the term has a constant meaning throughout this review.

Glossary

Adenocarcinoma cancer that begins in cells that line certain internal organs and that have glandular (secretory) properties

Aspiration removal of fluid from a lump, often a cyst, with a needle and a syringe

Benign not cancerous; does not invade nearby tissue or spread to other parts of the body

Biopsy a procedure used to remove cells or tissues in order to look at them under a microscope to check for signs of disease. When an entire tumour or lesion is removed, the procedure is called an excisional biopsy. When only a sample of tissue is removed, the procedure is called an incisional biopsy or core biopsy. When a sample of tissue or fluid is removed with a needle, the procedure is called a needle biopsy or fine-needle aspiration

Bone scan a technique to create images of bones on a computer screen or on film. A small amount of radioactive material is injected into a blood vessel and travels through the bloodstream. It collects in the bones and is detected by a scanner

Bronchitis inflammation of the bronchi

Bronchoscope a thin, lighted tube used to examine the inside of the trachea and bronchi, the air passages that lead into the lungs

Bronchoscopy a procedure in which a thin, lighted tube is inserted through the nose or mouth. This allows examination of the inside of the trachea and bronchi, which are the air passages that lead to the lung, as well as the lung itself. Bronchoscopy may be used to detect cancer or to perform some treatment procedures

Cancer a term for diseases in which abnormal cells divide without control. Cancer cells can invade nearby tissues and can spread through the bloodstream and lymphatic system to other parts of the body

Carcinogen any substance that causes cancer

Catheter a flexible tube used to deliver fluids into or withdraw fluids from the body

Chemotherapy treatment with anticancer drugs

Clinical trial a research study that tests how well new medical treatments or other interventions work in people. Each study is designed to test new methods of screening, prevention, diagnosis or treatment of a disease

Computed tomography scan a series of detailed pictures of areas inside the body; the pictures are created by a computer linked to an X-ray machine. Also called computed axial tomography (CAT) scan

Continuous hyperfractionated accelerated radiation therapy involves thrice daily treatment for 12 consecutive days; used for treatment of non-small-cell lung cancer

EMTREE EMBASE thesaurus (the basis of subject indexing in EMBASE)*

Epidermoid carcinoma a type of cancer in which the cells are flat and look like fish scales. Also called squamous cell carcinoma

External radiation radiation therapy that uses a machine to aim high-energy rays at the cancer. Also called external-beam radiation

Intravenous injected into a blood vessel

Large-cell carcinomas a group of lung cancers in which the abnormal cells are large

LIFETEST statistical procedure in Statistical Analysis Software (SAS)*

Lobe a portion of an organ such as the liver, lung, breast or brain

Lobectomy the removal of a lobe

Lymphatic system the tissues and organs that produce, store and carry white blood cells that fight infection and other diseases. This system

continued

Glossary *contd*

includes the bone marrow, spleen, thymus and lymph nodes, and a network of thin tubes that carry lymph and white blood cells. These tubes branch, like blood vessels, into all the tissues of the body

Magnetic resonance imaging a procedure in which a magnet linked to a computer is used to create detailed pictures of areas inside the body

Malignant cancerous; a growth with a tendency to invade and destroy nearby tissue and spread to other parts of the body

Mediastinoscopy a procedure in which a tube is inserted into the chest to view the organs in the area between the lungs and nearby lymph nodes. The tube is inserted through an incision above the breastbone. This procedure is usually used to get a tissue sample from the lymph nodes on the right side of the chest

Mediastinum the area between the lungs. The organs in this area include the heart and its large blood vessels, the trachea, the oesophagus, the bronchi and lymph nodes

Metastasis the spread of cancer from one part of the body to another. Tumours formed from cells that have spread are called 'secondary tumours' and contain cells that are like those in the original (primary) tumour. The plural is metastases

Non-small-cell lung cancer a group of lung cancers that includes squamous cell carcinoma, adenocarcinoma and large-cell carcinoma

Oat cell cancer a type of lung cancer in which the cells look like oats when viewed under a microscope. Also called small-cell lung cancer

PHREG statistical procedure in SAS*

Pneumonectomy an operation to remove an entire lung

Pneumonia an inflammatory infection that occurs in the lung

Prophylactic cranial irradiation radiation therapy given to the head to prevent cancer spreading to the brain. It is given to patients in complete remission from SCLC

Radiation therapy the use of high-energy radiation from X-rays, neutrons and other sources to kill cancer cells and shrink tumours. Radiation may come from a machine outside the body (external-beam radiation therapy) or from material called radioisotopes. Radioisotopes produce radiation and are placed in or near a tumour or near cancer cells. This type of radiation treatment is called internal-radiation therapy, implant radiation or brachytherapy. Systemic radiation therapy uses a radioactive substance, such as a radiolabelled monoclonal antibody, that circulates throughout the body. Also called radiotherapy

Radionuclide scanning a test that produces pictures (scans) of internal parts of the body. The person is given an injection or swallows a small amount of radioactive material. A machine called a scanner then measures the radioactivity in certain organs

Resection surgical removal of part of an organ

Respiratory system the organs that are involved in breathing. These include the nose, throat, larynx, trachea, bronchi and lungs

Side-effects problems that occur when treatment affects healthy cells. Common side-effects of cancer treatment are fatigue, nausea, vomiting, decreased blood cell counts, hair loss and mouth sores

Small-cell lung cancer a type of lung cancer in which the cells appear small and round when viewed under the microscope. Also called oat cell lung cancer

Sputum mucus coughed up from the lungs

Squamous cell carcinoma cancer that begins in squamous cells, which are thin, flat cells resembling fish scales. Squamous cells are found in the tissue that forms the surface of the skin, the lining of the hollow organs of the body, and the passages of the respiratory and digestive tracts. Also called epidermoid carcinoma

Thoracentesis removal of fluid from the pleural cavity through a needle inserted between the ribs

Thoracotomy an operation to open the chest

Tumour an abnormal mass of tissue that results from excessive cell division. Tumours perform no useful body function. They may be either benign (not cancerous) or malignant (cancerous)

X-ray high-energy radiation used in low doses to diagnose diseases and in high doses to treat cancer

* Used only in tables or appendices

List of abbreviations

AJCC	American Joint Commission for Cancer*
ALT	alanine transaminase*
AST	aspartate transaminase*
AUC	area under the curve*
BNF	British National Formulary
BSA	body surface area*
BSC	best supportive care
BUN	blood urea nitrogen*
CALGB	Cancer and Leukemia Group B*
CAV	cyclophosphamide, doxorubicin and vincristine*
CBDCA	carboplatin
CDDP	cisplatin
CEA	cost-effectiveness analysis*
CHART	continuous hyperfractionated accelerated radiation therapy*
CI	confidence interval*
COIN	Clinical Oncology Information Network
COSTART	Coding Symbols for Thesaurus of Adverse Reaction Terms*
CRC	Cancer Research Campaign
CRD	Centre for Reviews and Dissemination
CT scan	computed tomography scan*
DM	Deutschmark*
DOC	docetaxel
DRG	diagnosis-related group*
ECG	electrocardiogram*
ECOG	Eastern Cooperative Oncology Group*
ELVIS	Elderly Lung Cancer Vinorelbine Italian Study
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ	EORTC Quality of Life Questionnaire
EORTC QLQ-C30-LC13	lung cancer-specific quality of life instruments, covering both generic quality of life in patients with cancer and aspects specific to lung cancer
EPI	epirubicin
FACT-L	Functional Assessment of Cancer Therapy – Lung
FBC	full blood count*
FF	French francs*
5FU	5-fluorouracil
G-CSF	granulocyte colony-stimulating factor*
GEM	gemcitabine
GP	general practitioner
Hb	haemoglobin*
HRQoL	health-related quality of life
ICER	incremental cost-effectiveness ratio
IFOS	ifosfamide
IP	inpatient*
ITT	intention to treat
IUAC	International Union Against Cancer
iv	intravenous*
LCSS	Lung Cancer Symptom Scale
LON	lonidamine
LV	leucovorin
LVEF	left ventricular ejection fraction*
LYG	life-year(s) gained
LYS	life-year(s) saved
MCA	Medicines Control Agency
MER	merbarone
MeSH	Medical Subject Heading*
Mesna	[sodium 2-]mercaptoethane-sulphonate*
MIC	mitomycin, ifosfamide and cisplatin (chemotherapy combination)
MITO	mitomycin
MRC	Medical Research Council*
MRI	magnetic resonance imaging*

continued

List of abbreviations *contd*

MUGA	multiple gated acquisition (blood pool)*
MVP	mitomycin, vinblastine and cisplatin (chemotherapy combination)
NA	not applicable*
NCI	National Cancer Institute (USA)*
NCIC-CTC	National Cancer Institute of Canada – Common Toxicity Criteria*
NED	no evidence of disease*
NeLH	National electronic Library for Health
NICE	National Institute for Clinical Excellence (UK)
NS	not significant*
NSCLC	non-small-cell lung cancer
OP	outpatient*
PAX	paclitaxel
PCI	prophylactic cranial irradiation*
PE	cisplatin and etoposide chemotherapy combination drugs
PIR	piroxantrone
POHEM	Population Health Model (Canada)
PRBC	packed red blood cells*
PS	performance status*
Ptas	pesetas*
QALY	quality-adjusted life-year
QoL	quality of life
Q-TWiST	quality-adjusted time without symptoms or toxicity
RCT	randomised controlled trial
RSCL	Rotterdam Symptom Checklist
RTOG	Radiation Therapy Oncology Group*
SAS	Statistical Analysis Software*
SCLC	small-cell lung cancer
SD	standard deviation*
SE	standard error*
SESLS	Southeast Scotland Lung Study
SHPIC	Scottish Health Purchasing Information Centre
Skr	Swedish krona*
SS14	Subset of commonly reported symptoms from the EORTC QLQ-C30 and LC13 scales
SWOG	South West Oncology Group
TNM	pathological staging scheme for tumours (T, primary tumour; N, regional nodes; M, metastatic)*
TOI	Trial Outcome Index*
TWiST	time without symptoms or toxicity
VBL	vinblastine
VDS	vindesine
VM-26	teniposide
VNB	vinorelbine
VP-16	etoposide
WBC	white blood cell count*

* Used only in tables or appendices

Executive summary

Background

The incidence of lung cancer is declining following a drop in smoking rates, but it is still the leading cause of death from cancer in England and Wales, with about 30,000 deaths a year. Survival rates for lung cancer are poor everywhere, but they appear to be better in the rest of the European Community and the USA than in the UK. Only about 5% of people with lung cancer survive for 5 years, and nearly all of these are cured by surgery after fortuitously early diagnosis. At present, only a small proportion of patients (probably about 5%) with non-small-cell lung cancer are being given chemotherapy. Some centres treat a greater proportion.

Objectives

This review examines the clinical effectiveness and cost-effectiveness of four of the newer drugs – vinorelbine, gemcitabine, paclitaxel and docetaxel – used for treating the most common type of lung cancer (non-small-cell lung cancer). The first three drugs are used for first-line treatment, but at present docetaxel is used only after first-line chemotherapy has failed.

Methods

This report was based on a systematic literature review and economic modelling, supplemented by cost data.

Results

Number and quality of studies

A reasonable number of randomised trials were found – three for docetaxel, six for gemcitabine, five for paclitaxel and 13 for vinorelbine. The quality of the trials was variable but good overall. There was a wide range of comparators. Some trials compared chemotherapy with best supportive care (BSC), which involves care that aims to control symptoms, with palliative radiotherapy if needed, but not to prolong life. Others compared the newer drugs against previous drugs or combinations.

Summary of benefits

The gains in duration of survival with the new drugs are modest – a few months – but worthwhile in a condition for which the untreated survival is only about 5 months. There are also gains in quality of life compared with BSC, because on balance the side-effects of some forms of chemotherapy have less effect on quality of life than the effects of uncontrolled spread of cancer.

Costs

The total cost to the NHS of using these new drugs in England and Wales might be about £10 million per annum, but is subject to a number of factors. There would be non-financial constraints on any increase in chemotherapy for the next few years, such as staffing; the number of patients choosing to have the newer forms of chemotherapy is not yet known; and the costs of the drugs may fall, for example, as generic forms appear.

Cost per life-year gained

The available data did not provide an entirely satisfactory basis for cost-effectiveness calculations. The main problem was the lack of direct comparisons of the new drugs. In order to strengthen the analysis, three different modelling approaches were used: pairwise comparisons using trial data; cost-minimisation analysis, as if all the new regimens were of equal efficacy; and cost-effectiveness analysis pooling the results of several trials with different comparators, giving indirect comparisons of the new drugs by using BSC as the common comparator. A number of different scenarios were explored through extensive sensitivity analysis in each model. Outcomes were expressed in incremental cost per life-year saved or incremental cost, versus BSC. There was insufficient evidence from which to derive cost per quality-adjusted life-year.

In first-line treatment, vinorelbine, gemcitabine, and the lower-dose paclitaxel plus cisplatin combinations generally performed well against BSC under a range of different scenarios and especially when given as a maximum of 3 cycles. Incremental cost per life-year gained (LYG) versus BSC varied depending on scenario, but baseline figures based on trial data and protocols were: single-agent vinorelbine, £2194 per LYG; vinorelbine plus cisplatin, £5206; single-agent

gemcitabine, £5690; gemcitabine plus cisplatin, £10,041; and paclitaxel plus cisplatin, £8537. In second-line chemotherapy, docetaxel gave a cost per LYG of £17,546, again well within the range usually accepted as cost-effective.

However, in routine care, the impact of therapy would be regularly reviewed, and continuation would depend on response, side-effects, patient choice and clinical judgement. Chemotherapy would be stopped in non-responders, making chemotherapy more cost-effective. A 'real-life' scenario in which 60% of patients receive only 1 or 2 cycles of chemotherapy gives much lower costs per LYG, with single-agent gemcitabine, single-agent vinorelbine, and paclitaxel plus platinum appearing to be cost-saving compared with BSC; the incremental cost of gemcitabine plus cisplatin would be £2478 per LYG, and of vinorelbine plus cisplatin, £2808.

At the very least, gains in duration of survival were achieved without diminution of quality of life (at best, they improved quality) and with relatively low incremental cost.

Comparisons among the individual drugs should be viewed with caution because they have had to be based on indirect comparisons.

Limitations of the analysis

Each of the three models had limitations. The cost-effectiveness estimates from the pairwise comparisons were based on single studies. The cost-minimisation analysis assumed that the regimens have equal efficacy in practice. The cost-effectiveness analysis had to be based on pooling data from individual trials.

The costs of BSC, inpatient stay and outpatient visits were from Scottish data. Median rather than mean data on duration of survival have been used in the analysis, because most of the trials reported only median data. Median survival and number of drug cycles were calculated by averaging across a number of studies, rather than being reliant on one particular study. The costs of the less expensive antiemetics cited in the trials were omitted. The use of more modern and costly antiemetics would have a modest detrimental effect on cost-effectiveness. In the absence of published data, an estimate was made of the cost of side-effects of chemotherapy, in particular hospital admissions, and applied to all the new regimens. In practice, admissions related to side-effects and their respective costs are likely to vary by regimen.

Conclusions

The new drugs for non-small-cell lung cancer extend life by only a few months compared with BSC, but appear to do so without net loss in quality of life and at a cost per LYG that is much lower than for many other NHS activities. Depending on assumptions used, these new drugs range from being cost-effective, as conventionally accepted, to being cost-saving.

Implications of the newer drugs

One of the present constraints on chemotherapy is availability of inpatient beds. The advent of newer and gentler forms of chemotherapy given on an outpatient basis would not only overcome this, but it would allow more patients to be treated. This might apply particularly to older patients. The treatment of more patients would increase work-load for oncologists, cancer nurses and pharmacists. The Government has already announced increased expenditure on staff for cancer care. The previously pessimistic attitudes to chemotherapy in non-small-cell lung cancer are changing in the wake of the newer agents, and this shift is likely to increase referral.

Need for further research

Recent advances in chemotherapy are welcome, but their effects remain small for patients with non-small-cell lung cancer. Much more research is needed into better drugs, better combinations, new ways of assessing the likelihood of response and especially direct comparisons between the new regimens. This research would be aided by having a greater proportion of patients involved in trials, but there will be infrastructure implications of increased participation.

Chapter 1

Aim of the review

The aim of this review is to examine the clinical effectiveness of four of the newer drugs for lung cancer, taking into account their effect on both duration and quality of life (QoL), and to estimate their cost-effectiveness compared with other uses of resources. So far, these drugs are used only in the commoner type of lung cancer, which is known as non-small-cell lung cancer and which makes up around 80% of lung cancer. This report therefore does not address the treatment of small-cell lung cancer.

Chapter 2
Background

Description of underlying health problem

Despite reductions over recent decades, related to a drop in smoking rates, lung cancer is still the leading cause of death from cancer in England and Wales, and the third most common cause of death. There are about 29,000 deaths each year, of which 18,000 are in men and 11,000 in women. The outlook after diagnosis is poor, because unfortunately only a small proportion (about 10%) of patients are diagnosed in time for cure by surgery, and other forms of treatment are less successful. In England and Wales during 1986–90, 80% of patients died within a year of diagnosis and only 5% survived for 5 years. While 1-year survival rates improved by 2% between 1971–75 and 1986–90, there was no change in 5-year survival rates.[1]

However, survival rates, though still poor, appear to be better in the rest of the European Community than in England and Wales.[2] The 5-year survival rate is about 4% better in men (i.e. 9% vs 5%) and 5% better in women when the results in England and Wales are compared with the European average.[3] This suggests that there is scope for improvement in the UK. There is no evidence that the disease is more aggressive, or diagnosed later, in England and Wales than in other European countries, and it has been pointed out that the survival in best supportive care (BSC) arms of trials in the UK is as good as, or better than, that seen in the BSC arms in trials in Europe or North America (Thatcher N, Christie Hospital, Manchester: personal communication, 2001). In such a common cancer, even a small difference in survival rates would save many lives. For example, a 4% difference in 5-year survival (the difference between England and Wales vs the rest of the European Union) would equate to around 1200 lives saved per year.

Even within England, there are differences in survival rates,[4] with some areas having around 8% 5-year survival (e.g. Croydon, Southampton and South West Hampshire) and others having 2% (e.g. Rotherham and North Cumbria), as shown in *Figure 1*. These rates are based on cancer registrations over 4 years, and their confidence intervals of about 6% to 10% and 1% to 3%, respectively, are separate enough for us to be reasonably sure that the differences are real.

Incidence and prevalence

The incidence of lung cancer rises steeply with age (*Table 1*). Hence, 51% of men and 38% of women with lung cancer are over age 75 years. There are also social class differences, reflecting smoking prevalence. The total burden of disease has been falling due to a decline in the incidence in men; however, there has been an increase in the incidence in women. *Table 2* shows both social class and temporal trends.[5]

There are two main types of lung cancer, which are known as small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC). The latter is much commoner, comprising about 84% of all lung cancers.

Current service provision

Chemotherapy for NSCLC

Only a small proportion of patients in the UK with NSCLC have been receiving chemotherapy. Consultation with our panel of advisers suggests that, in England as a whole, only about 5% of people with this type of lung cancer are given chemotherapy, although the figure is higher in some areas. The highest estimate was 20%. In Wales, the report of the Welsh Thoracic Society found that 12% of patients diagnosed with lung cancer in 1996 were given chemotherapy.[6]

There seems to have been a widespread belief that chemotherapy for lung cancer was toxic and ineffective. A survey of clinicians who treated lung cancer in the UK found little support for chemotherapy.[7] This may reflect the published literature in previous years. However, more optimistic assessments have appeared in recent years. A meta-analysis published in 1993[8] concluded that there were small gains in survival and in QoL. An editorial appearing in 1994[9] noted that lung cancer remained "the Cinderella of cancer medicine" but that chemotherapy provided useful relief of

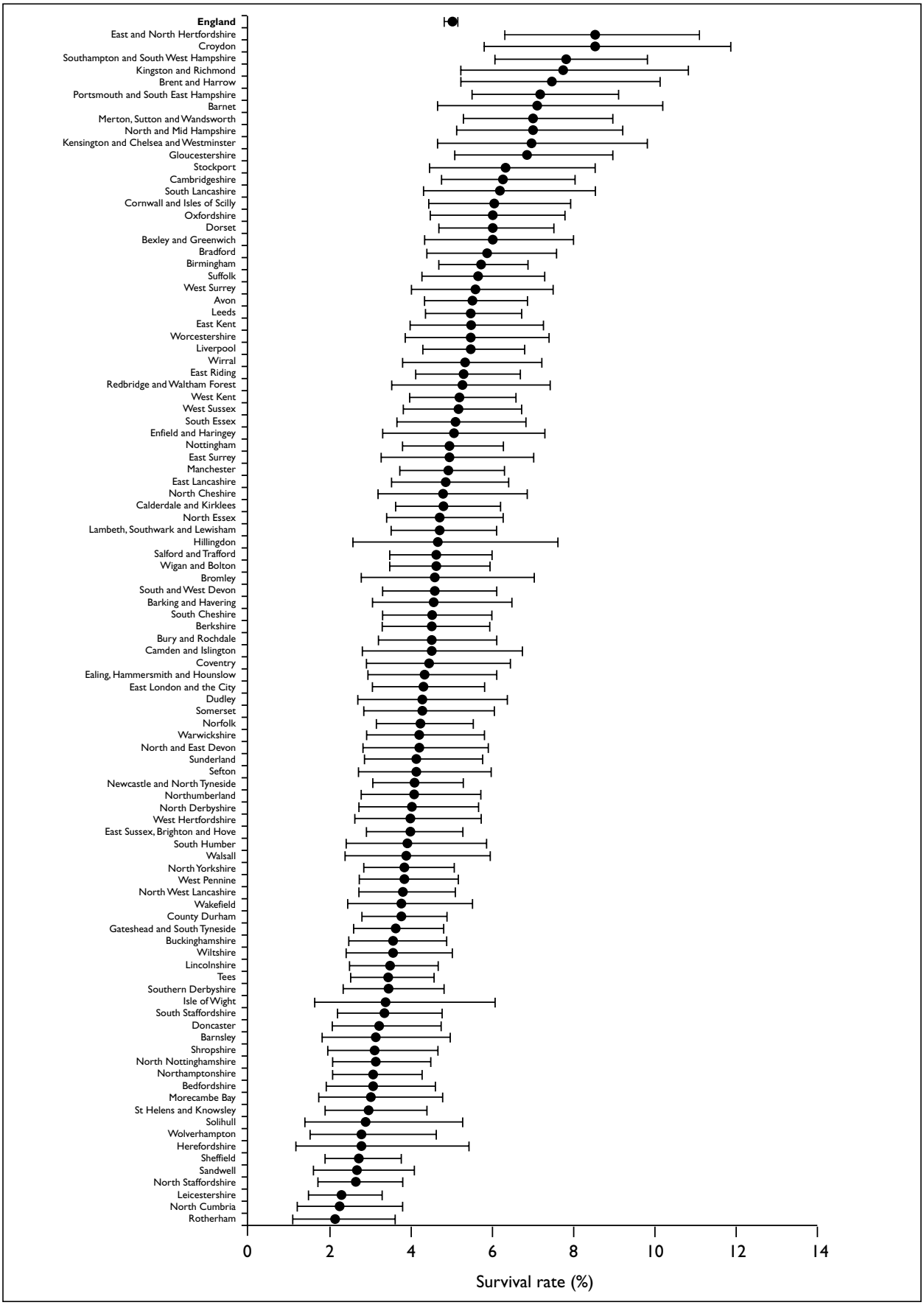


FIGURE 1 *Lung cancer survival: 5-year survival rates of persons aged 15–99 years who were diagnosed with lung cancer during the period 1991–1993 (source: Office of National Statistics)*

TABLE 1 *Incidence of lung cancer by age and sex*

	Registrations in 1996 per 100,000 population in England and Wales, by age (years)				
	15–44	**45–54**	**55–64**	**65–74**	**75 and over**
Men	2	41	154	416	627
Women	2	28	78	201	190

TABLE 2 *Incidence of lung cancer by social class and over time*

Social group	**Lung cancer incidence per 100,000 population during specified years[5]**		
	1970–1972	**1979–1983**	**1991–1993**
I	41	26	17
II	52	39	24
III (NM)	63	47	34
III (M)	90	72	54
IV	93	76	52
V	109	108	82
All groups	73	60	39

NM, non-manual; M, manual

symptoms such as malaise, breathlessness, cough and pain.[10] Another review in 1994[11] concluded that there was an increase in survival of a few months (4–7 months) but expressed doubts about the balance between the extra survival and the "toxicity and impaired quality of life", although this review did not produce any evidence on the QoL. One rather better review that did present such evidence[12] concluded that patients did benefit in terms of symptom control.

The large meta-analysis by the Non-small Cell Lung Cancer Collaborative Group,[13] published in the *BMJ* in 1995 but also as a Cochrane review last updated in February 2000, concluded that pessimism at the start of the review had changed to cautious optimism by the end. However, this well-conducted review may also explain part of the problem. The older chemotherapeutic regimens, using alkylating agents, probably did more harm than good. In the meta-analysis report, a Forrest plot of results, with the diamonds showing the aggregated results of trials, shows that the diamonds moved to the left over the period from 1970 to 1988, from a position where chemotherapy probably did more harm than good (the confidence interval did overlap with no effect) to a position where there was definite evidence of modest benefit mainly due to the advent of the platinum drugs.

Two other problems may have led to doubts about the value of chemotherapy. Firstly, some trials use response (defined as a shrinking in size, or disappearance of the tumour) as the outcome measure, and it is well known that response may not correlate with survival. However, the converse is also true – results from the UK gemcitabine study showed that, although only 20% of patients had a definite tumour size response, 50% had improvement in symptoms.[14]

Secondly, some commentators have expressed concern about the cost of new cancer drugs. However, the issue here is not simply the cost of the drug in isolation, but the total cost of care of patients with and without the drug. An expensive new drug may have its costs offset by reduced hospital stays (though it may be difficult in practice to realise the savings from these, because it would involve closing beds[15]), and what we need to know is the relative cost-effectiveness, preferably in terms of a common currency such as the cost per quality-adjusted life-year (QALY).

The Clinical Oncology Information Network (COIN) guidelines concluded that there was a place for combination therapy for NSCLC for selected patients, but recommended that chemotherapy should normally be given as part of a trial. The guidelines used evidence published up until

mid-1998, and were written when there was "strong evidence of only a modest effect on survival, with no clear evidence of benefit in terms of quality of life".[16] Other authors writing more recently suggested that the guidelines did not reflect the most recent advances and would not dispel "the nihilism surrounding lung cancer in the United Kingdom".[17]

The June 1998 report from the Centre for Reviews and Dissemination (CRD) expressed awareness that the four "new generation" drugs were becoming available, but noted that no published randomised controlled trials (RCTs) were then available.[18] This report supersedes the CRD one.

Patients' views

As part of this review, we sought advice from the Roy Castle Lung Cancer Foundation, which among many other activities, runs monthly patient support groups and a helpline, and therefore has frequent contacts with patients and families. Despite being a recent development, the helpline receives about 100 calls a month. About 25–30% of these calls are from patients or families seeking advice on alternatives to BSC. The Foundation is aware that those who phone the helpline may not be typical of all people with lung cancer (who tend to be older and from less-affluent social groups, and whose survival is often short). The opinions of those who phone include the following.

- If there is a possibility that chemotherapy could be beneficial, either by extending life or by maintaining QoL, patients feel that they should have the option of receiving this therapy.
- Patients are aware – perhaps sometimes from their doctors, perhaps from media, including the Internet – that the four new drugs have some activity against NSCLC and therefore offer some hope in a condition for which outcomes are so poor.
- Faced with an incurable disease, worsening symptoms and a short life expectancy, sufferers do not feel that cost should be a factor in deciding treatment options.
- Inequity in access to therapies, which are available in the private sector and in other countries, is seen as unjust.
- Patients often have an impression that their doctors believe that lung cancer has such a poor outlook that referral to specialist oncology services is not worthwhile.

The Foundation itself notes that these comments are inevitably heavily influenced by the emotion of a desperate situation and that they do not take account of the realities of health service funding.

Description of the interventions considered in this review

Docetaxel

Docetaxel received authorisation in 1999 for the treatment of patients with locally advanced or metastatic NSCLC after failure of prior chemotherapy. As with most anticancer agents, normal cells may be affected with severe adverse events, but these are often predictable and manageable. Possible adverse effects include allergic reaction presenting as flushing, skin reactions, itching, chest tightness, back pain, difficulty breathing, fever or chills, swelling, weight gain, stomach upsets, alopecia, cardiac irregularities and tiredness. Administration is usually via infusion once every 3 weeks. Contraindications include severe allergic reaction, low white blood cell count due to bone marrow damage ('myelosuppression'), severe liver disease, pregnancy or breastfeeding. Docetaxel is manufactured as Taxotere by Rhône-Poulenc Rorer (West Malling, UK).

Paclitaxel

Paclitaxel has been licensed for NSCLC in the UK since 1998, to be given in combination with platinum drugs, but not yet as monotherapy. Adverse events may include allergic reactions, blood disorders, fever, unusual bleeding or unexplained bruising, heart problems, high or low blood pressure, numbness, joint or muscle pain, liver disorders, nausea, diarrhoea, sore mouth and tongue, hair loss, skin reactions and swelling at the injection site. Doses are given by infusion, along with steroids given prior to administration of paclitaxel to help prevent allergic reaction. Paclitaxel is manufactured as Taxol by Bristol-Myers Squibb Pharmaceuticals (Hounslow, UK).

Gemcitabine

Gemcitabine is licensed for palliative treatment of locally advanced or metastatic NSCLC, either alone or in combination. Gemcitabine is given by intravenous infusion once weekly for 3 weeks. It should be stopped if there is evidence of haemolytic anaemia. Adverse events may include myelosuppression, lethargy, flu-like symptoms, rashes, nausea and vomiting, and hair loss. Gemcitabine (Gemzar) is produced by Eli Lilly (Basingstoke, UK).

Vinorelbine

Introduced in the UK in 1997 for the treatment of advanced NSCLC, vinorelbine is currently administered by intravenous infusion on a weekly basis in outpatients, though an oral version has been developed. It can be given alone or in combination. Adverse effects include myelosuppression, nausea and vomiting, constipation, weakness, peripheral neuropathy, alopecia and injection site pain. Vinorelbine (Navelbine) is produced by Pierre Fabre Ltd (Winchester, UK).

One of the commonest effects of most drugs used in cancer chemotherapy is damage to the bone marrow (myelosuppression), which can cause the numbers of both red and white blood cells to fall, producing anaemia in the case of the former and neutropenia in the latter. Reduced numbers of white blood cells can reduce resistance to infection, which can lead to serious infection needing hospital admission for antibiotic treatment, or sometimes death.

Assessing new treatments in cancer chemotherapy

In some types of cancer, the results of treatment are very good, with most patients being cured. Unfortunately, lung cancer has been one of the cancers most resistant to chemotherapy, and hence there has been a search for drugs that are more effective or less toxic, or both. When comparing new drug treatments, it is necessary to ask whether the new drug has benefits such as:

- longer survival
- improved QoL, which will be determined partly by toxicity.

QoL is particularly important if chemotherapy has only modest effects on the duration of survival. As has been noted above, the new drugs still have many side-effects.

Best supportive care

New drugs can be compared with older ones or with 'best supportive care'. This term needs a little elaboration and can be difficult to define. It does not refer simply to 'terminal care' measures such as pain relief. It implies treatment that aims to relieve symptoms but that does not attempt to prolong life. It is essentially palliative care, but it should be noted that BSC may include radiotherapy and sometimes chemotherapy. Some symptoms of advancing lung cancer, such as breathlessness (perhaps due to partial blockage of an airway), are difficult to deal with and may respond to chemotherapy.

The cost of BSC varies from place to place due to a number of factors, which can be grouped as follows:

1. service factors, such as the availability of beds, either in hospital or hospice
2. practice factors, such as radiotherapy regimens and the use of palliative chemotherapy
3. patients and social factors, such as urban/rural distribution and deprivation.

Costs of BSC are not available from routinely collected hospital data, partly because these data are categorised by disease not treatment, partly because much of the care is provided in the community. In any case, for the purposes of cost-effectiveness analysis, we need information about the components of care and the ways in which a change in treatment, such as greater use of chemotherapy, would affect the components and the overall cost, across all sectors. Greater use of chemotherapy might require more outpatient visits and laboratory support but fewer bed days than BSC.

One of the advantages of breaking the cost into its main components is that NHS trusts and health authorities can assess their own costs for each component. Some costs will vary little among hospitals (e.g. the cost of a drug or nurse salaries) but others will vary (e.g. old hospitals or old equipment will have lower capital charges, staff mixes may vary, and clinical practice and social factors will vary).

Chapter 3

Effectiveness

Methods for reviewing effectiveness

The *a priori* methods used for the rapid review are outlined in the research protocol (appendix 1), which was sent for expert comments to members of the advisory panel for the review.

The methods outlined in the protocol are summarised below.

- The review would consider SCLC and NSCLC as being two separate conditions. However, at present, the new drugs are not used in the treatment of SCLC.
- Studies of chemotherapy as an addition to surgery or radiotherapy were not included.
- Other tumours such as mesothelioma of the lung were not included, and neither were lung metastases from primary tumours outside the lung.
- The four drugs considered could be used in combination or separately (though not all are licensed for single use in the UK), and because of the number of possible combinations and comparators, some analyses would have to be indirect, using common comparators.
- Given the widespread pessimism in the UK about any chemotherapy for lung cancer, the default used for comparison would be BSC, but there would also be examination of the new drug regimens versus the better of the older regimens, with the baseline comparator being platinum-containing combinations.

Sources of information, including databases searched along with key search terms, can be found in appendix 2.

Studies used include systematic reviews, RCTs, QoL studies and economic appraisals. Data on incidence were obtained from Cancer Registry data. Studies available only in abstract form were not used in the efficacy analysis.

Studies identified through the search strategy were assessed for inclusion. Titles and abstracts were reviewed independently by two reviewers. The full text of included studies was examined independently by at least two reviewers. Data extraction and quality assessment were carried out by one reviewer and checked by a second reviewer. Any differences in opinion were resolved through discussion.

RCTs were assessed using the Jadad scale (appendix 3),[19] as is standard in most reviews for the National Institute for Clinical Excellence (NICE). However, in RCTs of chemotherapy versus BSC, blinding is clearly impossible, so the maximum score possible for these studies will be 3 not 5, and this will be noted in the report . In such studies, a score of 2 or more was considered to be of good quality. Blinding will be difficult in some other comparisons; these studies will be judged against a maximum score of 5 but should be interpreted with care. The use of the Jadad scale is under review. Systematic reviews were assessed for quality using the NHS CRD criteria.[20] Economics studies were reviewed using the appended appraisal questions (appendix 4).

QoL information was sought from the literature, to aid in the calculation of QALYs, if appropriate. Unpublished data were sought, and experts were consulted. References cited in industry submissions, sent to NICE by the four manufacturers specified on pages 6–7, were used as a check on completeness of ascertainment of relevant trials.

Evidence on patients' views was obtained from the literature and from the Roy Castle Lung Cancer Foundation.

Costing information was gathered partly from published NHS data, scientific publications, partly from unpublished work from the Scottish Health Purchasing Information Centre (SHPIC), which costed the care of a group of patients with lung cancer, and partly from data from the Southeast Scotland Lung Cancer Group.

Searching showed that there were no published trials on docetaxel, gemcitabine, paclitaxel and vinorelbine for SCLC, and the review is on NSCLC only.

Included studies used several different abbreviations to refer to the same intervention, and consultation with experts indicated that there was no universal set of abbreviations. Throughout the

remainder of the review, the different interventions will be referred to using the following abbreviations: docetaxel (DOC), gemcitabine (GEM), paclitaxel (PAX), vinorelbine (VNB), cisplatin (CDDP), combination of cyclophosphamide, doxorubicin and vincristine (CAV), carboplatin (CBDCA), epirubicin (EPI), 5-fluorouracil (5FU), ifosfamide (IFOS), lonidamine (LON), leucovorin (LV), merbarone (MER), mitomycin (MITO), cisplatin and etoposide in combination drugs (PE), piroxantrone (PIR), vinblastine (VBL), vindesine (VDS), teniposide (VM-26) and etoposide (VP-16).

The **key questions** were as follows.

1. Is chemotherapy with the new drugs clinically effective and cost-effective compared with BSC?
2. Does treatment with the new drugs have advantages over best current chemotherapy, taken to be the platinum combinations (based on the Cochrane review), such as mitomycin, ifosfamide and cisplatin (MIC) as well as mitomycin, vinblastine and cisplatin (MVP)?
3. If so, should the new drugs be used as single agents, or in combination with platinum or each other?

Outcomes used were survival and QoL. Details on toxicity were collected, partly for their effect on QoL, partly as a guide to the costs of treatment. Many studies reported 'response to treatment', but this has a specialised meaning, indicating tumour size. Complete response indicates disappearance, and partial response indicates 50% diminution in size of tumour. However, these responses may not be accompanied by any change in survival; conversely, there may be improvements in QoL without any 'response' being seen. The 50% used for defining partial response seems to have been based on change detectable with confidence on X-ray examination, and lesser reductions may be useful, for example, in opening up a blocked airway. Responses may be useful when assessing if a drug has an effect and as a guide as to whether to continue therapy, but for analysing clinical benefits to patients, symptoms and survival are better guides to efficacy. For completeness, we report responses when given in trial results, but they have not been used for comparative evaluations.

Docetaxel

Quantity and quality of research for docetaxel in NSCLC

DOC is at present licensed only for second-line treatment when other drugs or combinations have failed, but one trial examined its use in first-line therapy. Three RCTs met the inclusion criteria for the review and are shown in *Table 3* and appendix 5.[21–23]

Study quality

The three published RCTs had Jadad quality scores of 2. For two of the RCTs, this score suggested they were of reasonably good quality, due to the fact that the studies compared DOC with BSC, preventing blinding.[21,22] The reports of these two studies stated that they were randomised, although there was no description of the methods used, and provided an adequate description of withdrawals and dropouts.

The third study, which compared two doses of DOC with VNB or IFOS, scored poorly on the Jadad score because it was not blinded.[23] However, as the manufacturers pointed out, it would have been difficult to have blinding because the drugs were administered very differently: a 1-hour infusion of DOC once every 3 weeks, versus VNB infusions on days 1, 8 and 15 of each 3-week cycle. In addition, the patients given DOC also received prophylactic steroids. The report did not adequately describe the methods of randomisation, although descriptions of withdrawals and dropouts were given. Other concerns included the definition of intention-to-treat (ITT) analysis in one RCT[22] and the lack of statistical power for the analysis of specific drug comparisons in another RCT.[21] The main primary end-point used in the studies was survival, with secondary end-points being tumour response, duration of response, time to progression, QoL and toxicity.

Assessment of clinical effectiveness of docetaxel in NSCLC

One RCT assessed DOC versus BSC as first-line therapy (not currently a licensed indication in the UK),[22] while the other two RCTs compared differing regimens of DOC with BSC[21] or with VNB or IFOS as second-line treatment.[23]

Patient survival

Two RCTs reported an improvement in median survival rates for patients receiving DOC as first- and second-line therapy compared with those treated with BSC.[21,22] The improvement was significant (7.5 months vs 4.6 months, respectively) in the RCT of second-line therapy for the DOC dose of 75 mg/m^2.[21] In the RCT comparing 100-mg/m^2 DOC, 75-mg/m^2 DOC and VNB or IFOS as second-line therapy, there was no difference in median survival[23] (5.7 months vs 5.6 months with the 75-mg dose of DOC vs VNB

TABLE 3 *Summary of evidence of the effectiveness of docetaxel in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Shepherd *et al.*, 2000[21] Design: Not entirely clear, but probably Phase III. Multi-centre, randomised trial. ITT. Second-line therapy Intervention: DOC 100 mg/m^2 (49 patients), DOC 75 mg/m^2 (55 patients) every 3 weeks, BSC (100 patients) Patients: Stage IIIA/B or IV NSCLC, previously treated with CDDP-based chemotherapy (n = 204) Jadad quality score: 2/3	Complete response: DOC 100 mg/m^2, 0/49 (0%) DOC 75 mg/m^2, 0/55 (0%) Partial response: DOC 100 mg/m^2, 3/49 (6%) DOC 75 mg/m^2, 3/55 (5%)	Median survival time: BSC, 4.6 months (range, 3.7–6.0 months) DOC (both doses), 7 months (range, 5.5–9.0 months) (p = 0.047) DOC 100 mg/m^2, 5.9 months (p = 0.78) DOC 75 mg/m^2, 7.5 months (p = 0.01) One-year survival rate: BSC, 19% DOC (both doses), 29% DOC 100 mg/m^2, 19% DOC 75 mg/m^2, 37% BSC, 12%	QoL: All QoL parameters favoured DOC patients, and differences were statistically significant for pain and fatigue scales (p = 0.006 and 0.06, respectively). Significantly less tumour-related medications were used by DOC patients than BSC patients (p = 0.02)	Haematological toxicity: Toxicity was more frequent in patients on 100 mg/m^2 DOC compared with 75 mg/m^2 DOC, specifically grade 3 neutropenia, grade 3/4 anaemia, grade 3/4 febrile neutropenia Non-haematological toxicity: Toxicity (including nausea and vomiting) was observed in all groups. Diarrhoea occurred in DOC arm only. Infection occurred in more DOC patients than BSC patients
Roszkowski *et al.*, 2000[22] Design: Phase III, open-label, multicentre, randomised trial. ITT. First-line therapy Intervention: DOC 100 mg/m^2 (137 patients) every 3 weeks, BSC (70 patients) Patients: Stage IIIB or IV NSCLC (n = 207) Jadad quality score: 2/3	Complete response: DOC 100 mg/m^2, 2/137 (1.5%) Partial response: DOC 100 mg/m^2, 16/137 (11.7%) Response rate: ITT population, 13.1% (95% CI, 7.5% to 18.8%) Evaluable for response population, 19.6% (95% CI, 12.0% to 29.1%)	Median survival time: DOC, 6.0 months (95% CI, 5.0 to 8.0 months) BSC, 5.7 months (95% CI, 4.4 to 6.8 months) One-year survival rate: DOC, 25% BSC, 16% Two-year survival rate: DOC, 12% BSC, 0%	QoL: Emotional functioning was significantly in favour of the DOC arm in three of four methods of analysis: longitudinal mixed model (p = 0.01), pattern mixed model (p = 0.04) and worst score (p = 0.01), but not significant in AUC analysis. There was no difference between global health status/QoL and physical functioning scores for the two arms (p = NS). Using AUC analysis, there was significant improvement in nausea/vomiting (p = 0.04), pain (p < 0.0001) and dyspnoea (p = 0.02) in the DOC arm. Sensitivity analysis found in favour of the DOC arm for all measures of pain and dyspnoea, except nausea/vomiting. Except for diarrhoea, all dimensions favoured DOC over BSC (p = NS)	Haematological toxicity: BSC patients had a higher incidence of pulmonary events, neurocortical events and pain; DOC patients had a higher incidence of asthenia, infection and neurosensory events Non-haematological toxicity: Pulmonary events and pain were both more common for BSC patients than DOC patients, but nausea was higher for DOC patients. DOC patients used significantly less opiate and non-opiate analgesics than BSC patients (p < 0.0001 for both), tumour-related medications other than for pain, and palliative radiotherapy (p < 0.01 for both). DOC patients had a higher use of anti-infective drugs (p = 0.027)

continued

TABLE 3 contd *Summary of evidence of the effectiveness of docetaxel in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
Fossella *et al.*, 2000[23] Design: Phase III, open-label, multicentre, randomised trial. ITT. Second-line therapy Intervention: DOC 100 mg/m^2 (125 patients), DOC 75 mg/m^2 (125 patients), VNB or IFOS (123 patients), every 3 weeks Patients: Stage IIIB/IV NSCLC, previously treated with CDDP-based chemotherapy Jadad quality score: 2/5	Complete response: not reported Partial response: DOC 100 mg/m^2, 10.8% (95% CI, 6.1% to 18.1%) DOC 75 mg/m^2, 6.7% (95% CI, 3.1% to 13.1%) VNB or IFOS, 0.8% (95% CI, 0.0% to 5.3%) DOC 100 mg/m^2 vs VNB or IFOS (p = 0.001) DOC 75 mg/m^2 vs VNB or IFOS (p = 0.036) DOC 100 or 75 mg/m^2 vs VNB or IFOS (p = 0.002)	Median survival time: DOC 100 mg/m^2, 5.5 months DOC 75 mg/m^2, 5.7 months VNB or IFOS, 5.6 months One-year survival rate: DOC 100 mg/m^2, 21% (95% CI, 14% to 28%) DOC 75 mg/m^2, 32% (95% CI, 23% to 40%) VNB or IFOS, 19% (95% CI, 12% to 26%) DOC 100 mg/m^2 vs VNB or IFOS (χ^2 test, p = NS) DOC 75 mg/m^2 vs VNB or IFOS (χ^2 test, p = 0.025)	QoL: LCSS was used to evaluate QoL (reported in separate unpublished paper)	Haematological toxicity: The percentage of patients suffering from neutropenia (77% vs 54% vs 31%), febrile neutropenia (12% vs 8% vs 1%) and use of filgrastim (28% vs 7% vs 3%) differed significantly between DOC 100 mg/m^2 vs DOC 75 mg/m^2 vs VNB or IFOS, respectively Non-haematological toxicity: No significant variations between groups were found

ITT, intention to treat; CI, confidence interval; AUC, area under the curve; NS, not significant; LCSS, Lung Cancer Symptom Scale

or IFOS), but 1-year survival was 32% for the 75-mg DOC group versus 19% for the VNB/IFOS group (p = 0.025).

The 1-year survival rates were improved for the patients receiving DOC in all three studies. One trial reported 2-year survival rates, which were improved for patients receiving DOC compared with BSC (12% vs 0%, respectively) as first-line therapy.[22]

Quality of life

Two trials assessed QoL using the European Organisation for Research and Treatment of Cancer (EORTC) scale and Lung Cancer Symptom Scale (LCSS) scale.[21,22] Specific elements of patient QoL were improved for patients receiving DOC compared with BSC. One study found statistically significant improvements on pain (p = 0.006) and fatigue (p = 0.06) scales for patients receiving DOC second-line therapy, and there was less need for opiate analgesia (32% of DOC patients vs 49% of BSC patients) and palliative radiotherapy (26% vs 37%, respectively).[21] However, the main QoL results from this study have yet to be published. The second study reported significantly improved measures of emotional functioning ($p < 0.05$), nausea/vomiting (p = 0.04), pain ($p < 0.0001$) and dyspnoea (p = 0.02) resulting from DOC compared with BSC as first-line therapy, although there were no differences between global health status/QoL and physical functioning.[22] There was also less opiate analgesic use in the DOC group (41% vs 69%), less palliative radiotherapy (24% of DOC patients vs 41% of BSC patients) but more infections (58% of DOC patients had treatment vs 41% of BSC patients), and these figures support the more direct QoL measurements. However, there is a serious problem with differential reporting of QoL: questionnaires were missing in 41% of patients receiving BSC against 7% of patients receiving DOC, and this could have introduced considerable bias.

The RCT comparing DOC and VNB or IFOS as second-line therapy did not report QoL.[23]

Adverse effects

Adverse effects were reported in the three trials. One trial reported haematological toxic events for DOC, 100 mg/m^2 compared with 75 mg/m^2, with patients on the higher dose suffering more adverse toxic events.[21] Non-haematological toxicities were reported equally among both arms, although diarrhoea affected only patients receiving DOC. In the comparison of DOC versus BSC for second-line therapy, high toxic death rates among patients receiving DOC, 100 mg/m^2, were noted, necessitating a reduction in dose to 75 mg/m^2

of DOC for the second half of the study.[21] In the other RCT, patients receiving the different interventions suffered various haematological and non-haematological events, although patients on DOC used significantly less opiate and non-opiate analgesics ($p < 0.0001$) and tumour-related medications (other than for pain) and palliative radiotherapy ($p < 0.01$). Patients receiving DOC had higher rates of the use of anti-infective drugs ($p = 0.027$).[22] The RCT comparing 100-mg/m^2 DOC, 75-mg/m^2 DOC and VNB or IFOS found significantly higher reporting of neutropenia, febrile neutropenia and use of filgrastim with DOC than with VNB or IFOS.[23] There were no significant variations between groups for non-haematological adverse events.

Summary of the use of docetaxel in NSCLC

There were two trials of DOC in patients with NSCLC who have relapsed after previous chemotherapy, one comparing it with BSC, the other against VNB or IFOS.

Survival

Compared with BSC, DOC (75-mg dose) gave an extra 3 months of survival: an average of 7.5 months versus 4.6 months. After 1 year, 37% of patients in this DOC group were alive, compared with 19% of the BSC group. The 100-mg dose gave poorer results: an average survival of 5.9 months and 1-year survival of 19%. There was no difference in average survival compared with VNB/IFOS, but the proportion alive at 1 year was higher for the 75-mg DOC group.

Quality of life

QoL data for the two trials of DOC versus BSC are incomplete – either not yet fully reported in one study or with a large missing proportion in the other – but they do suggest improvement in the DOC groups; this is supported by evidence of reduced needs for opiate analgesia and palliative radiotherapy.

Gemcitabine

Quantity and quality of research for gemcitabine in NSCLC

Six RCTs met the inclusion criteria for the review and are shown in *Table 4* and appendix 6.[14,24–29]

Study quality

For all six RCTs, blinding was either impossible due to the comparator being BSC[14] or difficult because the comparators involved differing regimens.[24,26–29] As such, these studies could score only 3 on the Jadad quality score, and they are judged on this basis. Two of the RCTs were of good quality (Jadad score, 3), three were of reasonably good quality (Jadad score, 2), while one was of poor quality (Jadad score, 1). The good-quality RCTs adequately described their methods of randomisation and withdrawals/dropouts from the studies, but lacked description of any blinding in the studies. One RCT lacked any discussion of the statistical power of the study.[28] Although the RCTs of reasonably good quality were not blinded and did not describe the methods of randomisation, they did provide a description of the withdrawals and dropouts.[24,26–28] One of the RCTs of reasonably good quality lacked adequate discussion of the statistical power of the studies to detect significant effects on primary outcomes,[24] while two of the RCTs did not use ITT analysis for all comparisons.[24,27] The poor-quality RCT was not double-blind but mentioned that people were randomised to interventions, although no details were provided of the method used. In addition, no description of withdrawals and dropouts or the statistical power of the study to detect significant effects on the primary outcomes was provided.[29] Most studies lacked any reporting of statistical differences for adverse events.

The main primary end-points used in the studies were survival, objective tumour response rates, toxicity and QoL. Secondary end-points focused on QoL, disease-related symptoms, time-to-event parameters, survival, toxicity and objective tumour response rate.

Assessment of clinical effectiveness of gemcitabine in NSCLC

Several different combinations of interventions were assessed as comparators in the studies of GEM. Two RCTs[24,28] compared the GEM arm with the CDDP and VP-16 combination arm. One RCT looked at GEM and BSC as opposed to BSC alone.[14] Other RCTs assessed GEM with CDDP versus VP-16 with CDDP,[26] GEM and CDDP versus CDDP,[29] and GEM and CDDP versus MITO, IFOS and CDDP.[27]

Patient survival

All six RCTs assessed survival, whether median survival, survival to 1 or 2 years, or proportion of patients surviving. The two RCTs assessing GEM versus the CDDP and VP-16 combination reported decreases in median survival time for patients on GEM compared with those on the combination drugs, although the differences were

TABLE 4 *Summary of evidence of the effectiveness of gemcitabine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Anderson *et al.*, 2000[14] Design: Multicentre, randomised trial. ITT Intervention: GEM 1000 mg/m^2 with BSC (150 patients), BSC (150 patients) Patients: Symptomatic locally advanced or metastatic NSCLC (n = 300) Jadad quality score: 3/3	Complete response: Not reported Partial response: Not reported Overall response rate for GEM+BSC: 18.5% (95% CI, 13% to 26%) (not ITT)	Median survival time: GEM+BSC, 5.7 months (95% CI, 4.6 to 7.6 months) BSC, 5.9 months (95% CI, 5.0 to 7.9 months) (log-rank test, p = 0.84) Estimated 1-year survival rate: GEM+BSC, 25% BSC, 22% Estimated 2-year survival rate: GEM+BSC, 6% BSC, 7%	QoL: Data were assessed using both evaluable and unevaluable patients. At 2 months, improvements of ≥ 10% found for GEM+BSC patients were in emotional functioning, pain-symptom scale, chest pain, cough and fatigue, with BSC patients having an improvement in dyspnoea. Treatment differences in deterioration of ≥ 10% for GEM patients were in hair loss and role function, whereas for BSC patients they were in chest pain, shoulder pain and emotional functioning. At 4 months, improvements of ≥ 10% found in GEM patients were in chest pain, shoulder pain, emotional functioning, role domain, social domain and financial impact. Treatment differences in deterioration of ≥ 10% for GEM patients were in hair loss, whereas for BSC patients they were in social domain, pain-symptom scale and constipation SS14 (symptom scale): • Change from baseline to 2 months: GEM+BSC, –10% (improvement) BSC, +1% (deterioration) (2-sample t-test, p = 0.113) • Sustained (≥ 4 weeks) improvement of ≥ 25% in SS14 score: GEM+BSC, 22% BSC, 9% (Pearson's chi-squared test, p = 0.0014)	Grade 3–4 toxicity was low in the GEM arm. Neutropenia, infection, thrombocytopenia, nausea/vomiting, lethargy, rash and pulmonary toxicity were all reported in the GEM arm. More patients in the GEM arm, compared with BSC, had increased hair loss, ankle swelling and flu-like symptoms, but not skin rash

continued

TABLE 4 contd *Summary of evidence of the effectiveness of gemcitabine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Bokkel-Huinink *et al.*, 1999[24] (previous interim analysis reported by Manegold *et al.*, 1997[25]) Design: Phase II, multicentre, open-label, randomised study. Not ITT Intervention: GEM 1000 mg/m^2 (72 patients), CDDP 100 mg/m^2 with VP-16 100 mg/m^2 (75 patients) Patients: Stage IIIA (if inoperable), IIIB or IV NSCLC (n = 147) Jadad quality score: 2/5	Complete response: GEM, 0% CDDP+VP-16, 0% Partial response: GEM, 17.9% CDDP+VP-16, 15.3% (protocol qualified) (p = NS)	Median survival time: GEM, 6.6 months (95% CI, 4.9 to 7.3 months) CDDP+VP-16, 7.6 months (95% CI, 5.4 to 9.3 months) One-year survival rate: GEM, 26% CDDP+VP-16, 24% (p = NS)	QoL (physical, role, cognitive, emotional and social): No statistically significant difference between the two treatment arms was evident in change from baseline scores over 6-cycle treatment period (p > 0.05). There was no statistically significant difference in change from mean baseline scores for global QoL either between or within treatment arms (p > 0.05)	Haematological toxicity: CDDP+VP-16 patients had a higher incidence of grade 3 and 4 neutropenia (p < 0.001), and a significantly higher incidence of grade 3 and 4 thrombocytopenia (p = 0.003). GEM patients had a higher incidence of grade 3 and 4 anaemia (p = NS) Non-haematological toxicity: GEM patients had a higher incidence of grade 4 hair loss, grade 3 and 4 nausea/vomiting, grade 3 infection and grade 4 diarrhoea. CDDP+VP-16 patients had a higher incidence of grade 3 pulmonary events. The incidence of grade 4 infection was equal in both arms
Cardenal *et al.*, 1999[26] Design: Phase III, multicentre, randomised trial. ITT Intervention: GEM 1250 mg/m^2 with CDDP 100 mg/m^2 (69 patients) every 21 days, VP-16 100 mg/m^2 with CDDP 100 mg/m^2 (66 patients) every 21 days Patients: Stage IIIB or IV NSCLC (n = 135) Jadad quality score: 2/5	Complete response: GEM, 0% VP-16, 0% Partial response: GEM, 28 (40.6%; 95% CI, 29% to 53%) VP-16, 14 (21.9%; 95% CI, 13% to 34%) (p = 0.02)	Estimated median survival time: GEM, 8.7 months (95% CI, 7.7 to 10.2 months) VP-16 arm, 7.2 months (95% CI, 6.1 to 9.8 months) (p = 0.02) One-year survival probability: GEM, 32% VP-16, 26% (p = 0.19)	QoL: No clinically significant differences were evident in change from baseline within treatment arm or between treatment arms in functional domains or global QoL. Both groups saw significant improvement in pain, insomnia, cough, haemoptysis, chest pain and shoulder pain. There was no improvement in dyspnoea and fatigue for either arm. Peripheral neuropathy did not worsen in either arm. Both arms had significant worsening of nausea and alopecia. The only statistically significant difference between treatment arms in change from baseline was for alopecia, which was worse for the VP-16 arm	Haematological toxicity: The main toxic event was myelosuppression. Grade 3 and 4 neutropenia and febrile neutropenia were more pronounced in the VP-16 arm. Grade 4 neutropenia was twice as frequent in the VP-16 arm in comparison with GEM arm (p < 0.001). The GEM arm had a higher incidence of grade 3 anaemia, grade 3 neutropenia, grade 3 and 4 thrombocytopenia, PRBC transfusion and toxic death. The VP-16 arm had a higher incidence of grade 4 anaemia, grade 4 neutropenia, febrile neutropenia and platelet transfusion Non-haematological toxicity: The GEM arm had a higher incidence of grade 3 nausea/vomiting, grade 4 infection and grade 4 dyspnoea. VP-16 was associated with a greater incidence of grade 4 nausea/vomiting, grade 3 infection, grade 3 alopecia and grade 3 paresthesias. Grade 3 and 4 haemorrhage, fever, grade 3 dyspnoea, grade 4 alopecia and grade 4 paresthesias were of equal incidence in both arms

continued

TABLE 4 contd *Summary of evidence of the effectiveness of gemcitabine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Crino *et al*., 1999[27] Design: Phase III, multicentre, randomised trial. ITT analysis used for efficacy and toxicity, not for QoL data Intervention: GEM 1000 mg/m^2 with CDDP 100 mg/m^2 (155 patients) every 28 days; MITO 6 mg/m^2, IFOS 3000 mg/m^2 with CDDP 100 mg/m^2 (152 patients) every 28 days Patients: Stage IIIB or IV NSCLC (n = 307) Jadad quality score: 2/5	Complete response: GEM+CDDP, 1% Triple combination, 1% Partial response: GEM+CDDP, 37% Triple combination, 25% Overall response rate: GEM+CDDP, 38% (95% CI, 31% to 46%) Triple combination, 26% (95% CI, 19% to 33%) GEM+CDDP, significantly higher overall response rate (p = 0.029)	Overall median survival time: GEM+CDDP, 8.6 months Triple combination, 9.6 months No significant difference (Wilcoxon, p = 0.3393; log-rank test, p = 0.8771) One-year survival rate: GEM+CDDP, 33% Triple combination, 34%	QoL: Overall, there were no differences in changes in QoL between the two arms. The only differences between the two treatment arms for change from baseline were a worsening of alopecia in the triple combination arm and a greater improvement in chest pain in the GEM+CDDP arm (p < 0.05). Global QoL did not change significantly in either arm. Both arms noted a moderate decrease of physical functioning, also evidenced by worsening of fatigue and nausea/vomiting. Both arms noted an improvement in pain, insomnia and cough	Haematological toxicity: Myelosuppression was the main toxicity in both arms. The GEM+CDDP arm experienced a higher incidence of grade 3 neutropenia, grade 3 and 4 anaemia, and grade 3 and 4 thrombocytopenia. The triple combination arm experienced a higher incidence of grade 4 neutropenia. Incidence of grade 4 neutropenia was significantly higher in the GEM+CDDP arm (p < 0.001) Non-haematological toxicity: The triple combination arm experienced a higher incidence of grade 3 and 4 nausea/vomiting and grade 3 alopecia. Grade 4 alopecia was not observed at all. Grade 3 and 4 peripheral neuropathy, and moderate and severe dyspnoea were rare
Perng *et al*., 1997[28] Design: Phase II, randomised trial. ITT Intervention: GEM 1250 mg/m^2 (27 patients) every 28 days, CDDP 80 mg/m^2 with VP-16 80 mg/m^2 (26 patients) every 28 days Patients: Stage III (A or B) or IV NSCLC (n = 53) Jadad quality score: 3/5	Complete response: GEM, 0% CDDP+VP-16, 0% Partial response: GEM, 5 (19.2%; 95% CI, 8.3% to 30.1%) CDDP+VP-16, 5 (20.8%; 95% CI, 9.5% to 32.1%)	Median survival time: GEM, 37 weeks CDDP+VP-16, 48 weeks (log-rank test, Breslow test and Tarone-Ware, p = 0.65) One-year survival: Not reported	QoL: No data collected	Haematological toxicity: The CDDP+VP-16 arm had a higher incidence of grade 3 and 4 leucopenia, grade 3 thrombocytopenia and grade 3 anaemia. There were no incidences, in either arm, of grade 4 thrombocytopenia or grade 4 anaemia. Febrile neutropenia occurred in the CDDP+VP-16 arm Non-haematological toxicity: The CDDP+VP-16 arm had a higher incidence of grade 3 nausea/vomiting. No other grade 3 or 4 toxicities were reported in the GEM arm. However, the CDDP+VP-16 arm suffered grade 3 alopecia, neurological events and grade 4 diarrhoea

continued

TABLE 4 contd *Summary of evidence of the effectiveness of gemcitabine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
Sandler *et al.*, 2000[29] Design: Phase III, multicentre, randomised trial. ITT Intervention: GEM 1000 mg/m^2 with CDDP 100 mg/m^2 (260 patients) every 28 days, CDDP 100 mg/m^2 (262 patients) every 28 days Patients: Stage IIIA or IIIB or IV NSCLC (n = 522) Jadad quality score: 1/5	Complete response: GEM+CDDP, 1.2% CDDP, 0.4% Partial response: GEM+CDDP, 29% CDDP, 11% Overall response rate: GEM+CDDP, 30.4% CDDP, 11.1% (p < 0.0001)	Estimated median survival time: GEM+CDDP, 9.1 months (95% CI, 8.3 to 10.6 months) CDDP, 7.6 months (95% CI, 6.5 to 8.2 months) (log-rank test, p = 0.004) Estimated 1-year survival rate: GEM+CDDP, 39% CDDP, 28%	QoL: Baseline and median changes at last observation for each patient were not different between treatment arms. Patients in both arms noted a decrease in physical, functional well-being and total FACT-L scores, but no difference in other subscales; the changes were not statistically significant. Both arms noted a decrease in HRQoL, but there were no differences between arms	Haematological toxicity: The GEM+CDDP arm experienced a higher incidence of grade 3 and 4 anaemia, thrombocytopenia and granulocytopenia. The GEM+CDDP arm also had a higher incidence of hospitalisations for febrile neutropenia Non-haematological toxicity: There was no significant difference between arms. Both arms experienced grade 3 and 4 nausea, vomiting, pulmonary dyspnoea and neuromotor events. The GEM+CDDP arm had a higher incidence of malaise and asthenia

SS14, subset of commonly reported symptoms from the EORTC QLQ-C30 and LC13 scales; PRBC, packed red blood cells; FACT-L, Functional Assessment of Cancer Therapy – Lung; HRQoL, health-related quality of life

not statistically significant.[24,28] The 1-year survival probabilities and survival rates were improved for patients on GEM compared with patients on CDDP and VP-16 in one RCT, but again the differences were not statistically significant.[24] One RCT comparing GEM and BSC versus BSC alone found no difference in median survival or 1- and 2-year survival rates between patient groups, although this may have been related to a low performance threshold (i.e. less fit patients) for admission to the trial.[14] In addition, the survival with BSC was 6 months, which is longer than usual. The comparison of GEM and CDDP versus MITO, IFOS and CDDP showed a non-significant improvement in median survival for patients receiving MITO, IFOS and CDDP, and comparable 1-year survival rates for both arms.[27] The RCTs comparing GEM plus CDDP versus VP-16 plus CDDP,[26] and GEM plus CDDP versus CDDP[29] showed statistically significant improvements in median survival (p = 0.02 and 0.004, respectively), both of 1.5 months, for patients on GEM with respect to their comparators. In addition, 1-year survival rates favoured the patients on GEM plus CDDP versus their comparators, though not significantly,[26] or significance was not stated.[29]

Quality of life

Of the six RCTs of GEM, five reported QoL values.[14,24,26,27,29] The QoL measures used varied, with four RCTs[14,24,26,27] using the EORTC Quality of Life Questionnaire (QLQ)-C30 and lung cancer-specific LC13 measurement scales, one RCT using the SS14-derived symptom scale,[14] and one RCT using the Functional Assessment of Cancer Therapy – Lung (FACT-L) health-related QoL measure.[29]

In the RCT comparing GEM and BSC with BSC alone, the SS14 scale showed an improvement (–10%) in the GEM and BSC arm, with a deterioration (+1%) in the BSC arm between baseline and 2 months (p = 0.113).[14] A sustained improvement (≥ 25% for ≥ 4 weeks) in the SS14 score occurred significantly more frequently in the GEM and BSC arm (22%) compared with the BSC arm (9%) (p = 0.0014). At 2 months, the EORTC QLQ-C30 and LC13 measures showed an improvement greater than 10% in favour of GEM and BSC in five domains, compared with one domain for BSC alone, as well as a deterioration greater than 10% in two domains for the GEM plus BSC arm, compared with three domains for the BSC arm. At 4 months, all six domains

that showed improvements greater than 10% were in favour of the GEM and BSC arm, while one and three domains showed a 10% or more deterioration in the GEM plus BSC arm and the BSC arm, respectively.

The RCT comparing GEM and CDDP versus MITO, IFOS and CDDP reported no overall difference in QoL changes between the arms.[27] The only difference between the arms for change from baseline was a worsening in alopecia in the triple combination arm and a greater improvement in chest pain in the GEM and CDDP arm ($p < 0.05$). The study assessing GEM plus CDDP against VP-16 plus CDDP reported no overall significant differences between treatment arms in changes in QoL from baseline.[26,29] Both arms had a significant improvement in pain, insomnia, cough, haemoptysis, chest pain and shoulder pain, and worsening of nausea and alopecia. Only alopecia differed significantly between treatment arms, worsening in the VP-16 arm. The study comparing GEM against CDDP plus VP-16 reported no significant difference between arms for functional or global scales.[24] In the comparison of GEM and CDDP to CDDP alone,[29] no difference was reported between interventions on the FACT-L questionnaire.

Adverse effects

Toxic events included grade 3 and 4 anaemia, neutropenia, thrombocytopenia, hair loss, nausea, infection and diarrhoea. There were no marked differences in side-effects between GEM and older regimens.

Summary of the use of gemcitabine in NSCLC

Survival

Five trials compared GEM with older treatments: three showed no difference in survival and two showed gains of 1.5 months. One trial comparing GEM with BSC found no difference in survival, but this might be explained by a better than expected survival in the BSC group. This trial was designed with QoL as the main outcome.

Quality of life

QoL was improved by GEM compared with BSC alone, and this improvement was accompanied by a marked reduction in palliative radiotherapy: more patients receiving BSC (79%, vs 49% of the GEM group) had radiotherapy, and they had it much earlier (by 4 weeks with BSC, compared with 29 weeks in the GEM group). There were few differences in QoL with GEM compared with older drugs.

Paclitaxel

Quantity and quality of research for paclitaxel in NSCLC

Six RCTs met the inclusion criteria for the review and are shown in *Table 5* and appendix 7.[30–35]

Study quality

Two of the six published RCTs were of good quality (Jadad score, 2 or 3 out of 3),[32,35] while the other four RCTs were of relatively poor quality (Jadad score, 2 out of 5).[30,31,33,34] None of the studies were blinded. For the best-quality RCT, blinding was not possible because PAX and BSC were compared with BSC.[35] However, the trial's methods of randomisation and withdrawals and dropouts were described.[35] Four of the RCTs did not adequately describe the methods of randomisation.[30–33] One RCT failed to state whether the study analysis was an ITT analysis.[32] One of the poor-quality RCTs was an interim analysis,[33] with limited reporting of all end-points and poor definition of outcomes. One RCT did not adequately report QoL data.[34]

The main primary end-points in the RCTs were survival and response, with secondary end-points being QoL, toxicity, time to treatment failure, and survival as well as progression-free survival.

Assessment of clinical effectiveness of paclitaxel in NSCLC

Patient survival

Two of the six RCTs showed improved median survival for patients receiving PAX,[30,32] with one RCT reporting a significant difference of 2 months compared with BSC ($p = 0.037$)[32] and the other an improvement of 2 months when PAX plus CDDP was compared with CDDP plus VP-16.[30]

Another three RCTs found no significant differences in median survival between PAX combinations and their comparators.[31,34,35] One- and 2-year survival was improved for patients receiving PAX combinations in four RCTs assessing survival,[30–32,35] but this difference was statistically significant in only one trial.[32] A fifth RCT reported a shorter median survival time and a lower estimated survival rate for patients receiving PAX.[34] The sixth RCT did not report any survival data.[33]

Quality of life

Four RCTs reported QoL data.[30,32,34,35] Various QoL scales were used, including the FACT-L,[30] Rotterdam Symptom Checklist (RSCL),[32] and EOTRC QLQ-C30 and LC13 scales.[34,35] Comparison of changes in QoL showed some improve-

TABLE 5 *Summary of evidence of the effectiveness of paclitaxel in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Bonomi *et al.*, 2000[30] Design: Phase III, multicentre, randomised trial. Not ITT Intervention: VP-16 100 mg/m^2 with CDDP 75 mg/m^2 (193 patients), PAX 250 mg/m^2 with CDDP 75 mg/m^2 (191 patients), PAX 135 mg/m^2 with CDDP 75 mg/m^2 (190 patients), every 21 days Patients: Stage IIIB or IV NSCLC (n = 574) Jadad quality score: 2/5	Complete response: Not reported Partial response: Not reported Overall response rate (p based on comparison with VP-16+CDDP): VP-16+CDDP, 12.4% PAX 250 mg/m^2+ CDDP, 27.7% (p < 0.001) PAX 135 mg/m^2+ CDDP, 25.3% (p = 0.002) PAX 250 mg/m^2+ CDDP vs PAX 135 mg/m^2+CDDP (p = 0.246)	Median survival time: CDDP+VP-16, 7.6 months PAX 250 mg/m^2+ CDDP, 10 months PAX 135 mg/m^2+ CDDP, 9.5 months One-year survival rate: CDDP+VP-16, 31.8% PAX 250 mg/m^2+ CDDP, 40.3% PAX 135 mg/m^2+ CDDP, 37.4%	QoL: A higher percentage of PAX patients had improved QoL (p = 0.46). If missing data were included, then there was a significantly higher rate of improved QoL for PAX patients (p = 0.012)	Haematological toxicity: Toxicity was equally evident in all arms of the trial (p not stated). A high level of grade 4 leucopenia was found in the PAX 250-mg/m^2 arm and high level of grade 4 granulocytopenia in the PAX 135-mg/m^2 arm, whereas a low level of grade 4 thrombocytopenia was found in the PAX 135-mg/m^2 arm Non-haematological toxicity: Grade 3 toxicity was equivalently evident in all arms. The only significant difference (p = 0.026) was in grade 5 cardiac events, which were more evident in the PAX 250-mg/m^2 arm
Chang *et al.*, 1993[31] Design: Phase II, randomised study. Not ITT Intervention: PAX 250 mg/m^2 (25 patients), MER 1000 mg/m^2 (35 patients), PIR 150 mg/m^2 (44 patients), every 3 weeks Patients: Stage IV NSCLC (n = 113) Jadad quality score: 2/5	Complete response: PAX, 0 (0%) MER, 0 (0%) PIR, 1/44 (2.3%) Partial response: PAX, 5/24 (20.8%; 95% CI, 7% to 42%) MER, 2/35 (5.7%) PIR, 0% Significance not stated	Median survival time: PAX, 24.1 weeks MER, 19.9 weeks PIR, 29.3 weeks (p = NS) One-year survival rate: PAX, 41.7% (± 10%) MER, 21.6% (± 7%) PIR, 22.6% (± 7%) (p = NS)	QoL: No QoL data collected	Haematological toxicity: Toxicity was more frequent in the PAX arm compared with the other two arms, after which it was more frequent in the PIR arm than in the MER arm Non-haematological toxicity: Toxicities were more frequent in the PAX arm than the other two treatment arms, except for grade 3–4 nausea/vomiting and grade 3–4 neuromotor, neuropsychological, genitourinary and phlebitis events, which were greater in the MER treatment arm. The grade 5 toxicities were more evident in the PAX and MER arms
Ranson *et al.*, 2000[32] Design: Phase III, multicentre, randomised trial. ITT not specified Intervention: PAX 200 mg/m^2 with BSC (79 patients), BSC (78 patients), every 21 days Patients: Stage IIIB or IV NSCLC (n = 157) Jadad quality score: 2/3	Complete response: PAX+BSC, 1/76 (1%) Partial response: PAX+BSC, 11/76 (14%) Overall response: PAX+BSC, 12/76 (16%; 95% CI, 8% to 26%) Not statistically significant	Median survival time: PAX+BSC, 6.8 months (95% CI, 5.7 to 10.2 months) BSC, 4.8 months (95% CI, 3.7 to 6.8 months) (log-rank test, p = 0.037) One-year survival rate: PAX+BSC, 95% CI, 20% to 41% BSC, 95% CI, 18% to 39% PAX+BSC significantly associated with increased survival (p = 0.048) Hazard ratio, 0.68 (95% CI, 0.49 to 1.0)	QoL: Compared with BSC, there was improvement in the functional activity subscore favouring the PAX+BSC arm for dropouts (p = 0.043). For all other subscores, there was no statistically significant difference between arms	Haematological toxicity: Grade 3–4 adverse events were more evident for the PAX+BSC arm Non-haematological toxicity: Grade 3–4 adverse events were more evident for the PAX+BSC arm

continued

TABLE 5 contd *Summary of evidence of the effectiveness of paclitaxel in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Postmus *et al*., 1996[33] Design: Phase II, multicentre, randomised trial. Interim analysis. ITT not specified Intervention: CDDP 80 mg/m^2 with VM-26 100 mg/m^2 (38 patients), PAX 175 mg/m^2 with CDDP 80 mg/m^2 (35 patients) Patients: Stage IIIB or IV NSCLC (*n* = 73) Jadad quality score: 2/5	Complete response: CDDP+VM-26, 0% PAX+CDDP, 3% Partial response: CDDP+VM-26, 26% PAX+CDDP, 37% Overall response: Not reported Significance not stated	Survival: Not reported	QoL: No data collected	Haematological toxicity: Grade 3–4 toxicity was more evident in the CDDP+VM-26 arm Non- haematological toxicity: Toxicity was more evident in the CDDP+VM-26 arm, except for myalgia, and neurological events, which were more evident in the PAX+CDDP arm. Significance not stated
Gatzemeier *et al*., 2000[34] Design: Phase III, multicentre randomised trial. ITT Intervention: PAX 175 mg/m^2 with CDDP 80 mg/m^2 (207 patients), CDDP 100 mg/m^2 (207 patients), every 3 weeks Patients: Stage IIIB or IV NSCLC (*n* = 414) Jadad quality score: 2/5	Complete response: PAX+CDDP, 2% CDDP, 1% Partial response: PAX+CDDP, 25% CDDP, 17% Overall response: PAX+CDDP, 50/207 (24%; 95% CI, 18% to 31%) CDDP, 34/207 (16%; 95% CI, 12% to 22%)	Survival time: PAX+CDDP, 8.1 months (95% CI, 7.3 to 9.2 months) CDDP, 8.6 months (95% CI, 7.1 to 10.3 months) Estimated 1-year survival rate: PAX+CDDP, 30% CDDP, 36%	QoL: There was a significant difference from baseline to various treatment periods for the PAX+CDDP arm in the symptom scales of nausea and vomiting ($p < 0.0003$), appetite loss ($p < 0.02$) and constipation ($p < 0.032$). For the CDDP arm, the significant differences were for hair loss and peripheral neuropathy ($p < 0.0001$)	Haematological toxicity: Grade 3/4 adverse events were significantly higher in the PAX+CDDP arm for neutropenia and febrile neutropenia ($p < 0.05$) Non-haematological toxicity: There were no significant differences between the two arms
Giaccone *et al*., 1998[35] Design: Phase III, randomised trial. ITT Intervention: CDDP 80 mg/m^2 with VM-26 100 mg/m^2 (166 patients), PAX 175 mg/m^2 with CDDP 80 mg/m^2 (166 patients), every 3 weeks Patients: Locally advanced or metastatic NSCLC (*n* = 332) Jadad quality score: 3/5	Complete response: CDDP+VM-26, 1 (1%) PAX+CDDP, 2 (1%) Partial response: CDDP+VM-26, 44 (27%) PAX+CDDP, 61 (39%) Overall response rate: CDDP+VM-26, 28% (95% CI, 21% to 35%) PAX+CDDP, 41% (95% CI, 33% to 48%) ($p = 0.018$)	Median survival time: CDDP+VM-26, 9.9 months PAX+CDDP, 9.7 months ($p = 0.971$) One-year survival rate: CDDP+VM-26, 41% (95% CI, 33% to 49%) PAX+CDDP, 43% (95% CI, 25% to 51%) Two-year survival rate: CDDP+VM-26, 18% (95% CI, 10% to 26%) PAX+CDDP, 19% (95% CI, 12% to 26%)	QoL: For most of the functioning scales, there was a significant benefit in favour of the PAX+CDDP arm at 6-week assessment on the symptom scale; fatigue ($p = 0.006$) and appetite loss ($p < 0.001$) were significantly lower at 6 weeks for the PAX+CDDP arm. However, the benefit was no longer evident at the 12-week assessment. Greater peripheral neuropathy was found in PAX+CDDP patients at 12 weeks	Haematological toxicity: Grade 3–4 adverse events were significantly more evident in the CDDP+VM-26 arm, except for grade 4 neutropenia, which was more evident in the PAX+CDDP arm Non-haematological toxicity: Grade 3–4 infection incidence was significantly higher ($p = 0.02$) in the CDDP+VM-26 arm. Grade 3 myalgia and peripheral neurotoxicity were more evident in the PAX+CDDP arm. Grade 3 stomatitis and other adverse events were equally evident in both arms

ment for patients receiving PAX, although these changes were not usually significant. The RCT comparing PAX plus CDDP with CDDP plus VP-16 [30] reported an improvement in QoL, but the difference was not statistically significant. The RCT of PAX plus BSC versus BSC alone[32] showed a significant improvement in functional ability for patients receiving PAX plus BSC, but no other differences were significant. In contrast, the RCT comparing PAX plus CDDP with CDDP plus VM-26[35] found a significant improvement in most functioning scales. However, significant decreases in fatigue and appetite loss for the PAX plus CDDP arm at the 6-week assessment were not evident at 12 weeks. One RCT comparing PAX plus CDDP with CDDP alone did not adequately report QoL data.[34]

Adverse effects

Adverse events were reported by all six RCTs, although the significance of any differences was only reported in three RCTs.[30,34,35] Comparative frequency of haematological and non-haematological events varied depending on the comparators. Three RCTs, one comparing PAX with MER and PIR, one comparing PAX plus CDDP with CDDP alone, and another comparing PAX and BSC with BSC only, found severe adverse events (grade 3 or 4) were more frequent in the PAX arm of the study.[31,32,34] In contrast, two of the three RCTs comparing CDDP and VM-26 with PAX and CDDP showed severe adverse events to be higher in the CDDP and VM-26 arm.[33,35] The third RCT found haematological and non-haematological events to be equally reported in the three arms.[30] The main adverse events reported in the studies for PAX were thrombocytopenia, leucopenia, anaemia, alopecia and nausea/vomiting. There were more toxicity-related cardiac deaths in the PAX 250-mg/m^2 group ($p = 0.026$),[30] while patients receiving CDDP and VM-26 reported significantly higher rates of grade 3–4 infection.[35]

Summary of the use of paclitaxel in NSCLC

Survival

In the one trial comparing PAX with BSC, PAX improved survival by 2 months. In the RCTs comparing it with other drugs, results were mixed, with no clear advantage in survival.

Quality of life

Compared with BSC, QoL in patients receiving PAX was similar for most end-points and better for one end-point. Hence, the gain in survival was not at the cost of poorer QoL. Two trials comparing PAX with other drugs found little difference in QoL: one reported no difference in QoL, and the other found that patients on PAX were better at 6 weeks and worse at 12 weeks.

Vinorelbine

Quantity and quality of research for vinorelbine in NSCLC

Thirteen RCTs met the inclusion criteria for the review and are shown in *Table 6* and appendix 8.[36–48]

Study quality

Eight of the 13 published RCTs had scores on the Jadad Quality Scale of 3 out of 5,[36,38–43,48] four RCTs had scores of 2 out of 5,[44–47] and one RCT scored 1 out of 5.[37] Of the eight RCTs scoring 3 on the Jadad Quality Scale, seven were considered of reasonably good quality because they described their methods of randomisation, as well as providing a description of withdrawals and dropouts. However, no description of blinding was provided, although blinding appeared possible. The other RCT with a score of 3 on the Jadad Quality Scale was considered to be of good quality.[48] This study also adequately described its methods of randomisation and withdrawals/dropouts but failed to blind interventions. However, it compared VNB with BSC, which would be impossible to blind. It should be noted that this RCT was stopped early by investigators (who were blinded to results) due to the low enrolment rate. Other issues of methodological concern with these RCTs included selective reporting of patients' baseline characteristics for eligible patients only,[39] and poor reporting of QoL results[40] and haematological toxicities.[41] The four RCTs scoring 2 on the Jadad Quality Scale did not give an adequate description of the method of randomisation or the method of blinding, but provided descriptions of withdrawals and dropouts. For three RCTs, blinding was not possible due to their comparator being CDDP[44,47] or the study being an open-label trial.[46] While these studies were considered to be of reasonably good quality, the other RCT was considered of poor quality because blinding, although difficult, appeared possible.[45] Other methodological concerns included poor or partial reporting of results as ITT,[46,47] limited reporting of background characteristics of patients[47] and poor reporting of results.[44] The very poor quality RCT did not give a description of the method of randomisation or the method of blinding, if used, or a description of withdrawals and dropouts.[37]

TABLE 6 *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
Baldini *et al.*, 1998[36] Design: Phase II, multicentre, randomised study. ITT Intervention: CDDP 80 mg/m^2 with VDS 3 mg/m^2 and MITO 6 mg/m^2 every 28 days (49 patients), CDDP 80 mg/m^2 with IFOS 3 mg/m^2 and VNB 25 mg/m^2 every 21 days (48 patients), CBDCA 350 mg/m^2 with VNB 25 mg/m^2 every 28 days (43 patients) Patients: Stage IIIB or IV NSCLC (n = 140) Jadad quality score: 3/5	Complete response: CDDP+VDS+MITO, 0% CDDP+IFOS+VNB, 0% CBDCA+VNB, 0% Partial response: Not reported Overall response rate: CDDP+VDS+MITO, 14.3% (95% CI, 5.9% to 27.2%) CDDP+IFOS+VNB, 16.7% (95% CI, 7.4% to 30.2%) CBDCA+VNB, 14% (95% CI, 5.3% to 27.9%)	Median survival time: CDDP+MITO+VDS, 8.4 months CDDP+IFOS+VNB, 8.8 months CBDCA+VNB, 7.9 months (p not stated) One-year survival rate: CDDP+MITO+VDS, 18% CDDP+IFOS+VNB, 15% CBDCA+VNB, 16%	QoL: No QoL data	Haematological toxicity: The main adverse event was myelosuppression. Haematological toxicity was more frequent in the CDDP+VNB+IFOS arm, except for grade 4 leucopenia, grade 3 thrombocytopenia and grade 4 anaemia, which were more evident in the CBDCA+VNB arm. Grade 4 nephrotoxicity was reported in 2.1% of CDDP+VNB+IFOS patients only Non-haematological toxicity: Only grade 3 nausea/vomiting was reported, with patients in the CDDP+IFOS+VNB arm suffering the most, followed by the CDDP+VDC+MITO patients
Colleoni *et al.*, 1997[37] Design: Phase II, randomised trial. ITT not specified Intervention: CDDP 100 mg/m^2 with MITO 8 mg/m^2 and VNB 25 mg/m^2 every 21 days (26 patients), CBDCA 400 mg/m^2 with VNB 25 mg/m^2 every 21 days (26 patients) Patients: Stage IIIB and IV NSCLC (n = 52) Jadad quality score: 1/5	Complete response: CDDP+MITO+VNB, 0% CBDCA+VNB, 0% Partial response: CDDP+MITO+VNB, 11 (42%) CBDCA+VNB, 7 (27%) Overall response: CDDP+MITO+VNB, 11/26 (42%; 95% CI, 23% to 63%) CBDCA+VNB, 7/26 (27%; 95% CI, 12% to 48%)	Median survival time: CDDP+MITO+VNB, 9.9 months (range, 3–14 months) CBDCA+VNB, 8.8 months (range, 1–18 months) One-year survival: Not reported	QoL: No QoL data	Haematological toxicity: Toxicity was more evident in the CBDCA+VNB treatment arm Non-haematological toxicity: Toxicity was not very evident in either arm. The CBDCA+VNB arm suffered from grade 3–4 mucositis, and the CDDP+VNB+MITO arm suffered from grade 3-4 vomiting and renal toxicity only

continued

The main primary end-points reported in the studies were survival and tumour response. The secondary end-points were duration of tumour response, toxicity, QoL, time to event parameters and survival.

Assessment of clinical effectiveness of vinorelbine in NSCLC

The 13 RCTs assessed 12 different comparisons of VNB (usually in combination) with other interventions. Five RCTs compared different doses of VNB and/or different combinations.[36,37,41,43,44] Two RCTs used a form of crossover design,[38,42] although patients in one RCT changed interventions only when considered to be non-responders.[42] Different combinations of the VNB arm were used in the RCTs, ranging from VNB alone to VNB+CDDP, VNB+CBDCA, VNB+CBDCA+CDDP, VNB+MITO+CDDP, VNB+CDDP+IFOS, VNB+CDDP+IFOS+EPI and VNB+CDDP+CBDCA+VP-16.

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Colucci *et al.*, 1997[38] Design: Phase III, multicentre, prospective, randomised study. ITT Intervention: Two-step treatment – CDDP 100 mg/m^2 with VNB 25 mg/m^2 every 21 days for 3 cycles, followed by IFOS 2.5 g/m^2 and EPI 100 mg/m^2 every 21 days for 3 cycles (53 patients); IFOS 2.5 g/m^2 and EPI 100 mg/m^2 every 21 days for 3 cycles, followed by CDDP 100 mg/m^2 and VNB 25 mg/m^2 every 21 days (47 patients) Patients: Stage IIIA/B and IV NSCLC (n = 100) Jadad quality score: 3/5	Complete response: • Post-step 1: CDDP+VNB (IFOS+EPI), 2% IFOS+EPI (CDDP+VNB), 2% • Post-step 2: CDDP+VNB (IFOS+EPI), 0% IFOS+EPI (CDDP+VNB), 5% Partial response: • Post-step 1: CDDP+VNB (IFOS+EPI), 45% IFOS+EPI (CDDP+VNB), 19% • Post-step 2: CDDP+VNB (IFOS+EPI), 0% IFOS+EPI (CDDP+VNB), 21% Overall response rate: • Post-step 1: CDDP+VNB (IFOS+EPI), 47% (95% CI, 33% to 61%) IFOS+EPI (CDDP+VNB), 21% (95% CI, 11% to 35%) (p = 0.0112) • Post-step 2: CDDP+VNB (IFOS+EPI), 0% IFOS+EPI (CDDP+VNB), 26% (p = 0.037)	Median survival time: CDDP+VNB (IFOS+EPI), 9 months IFOS+EPI (CDDP+VNB), 7 months (p = 0.328) One-year survival: Not reported	QoL: No QoL data	Haematological toxicity: In steps 1 and 2, the IFOS+EPI (CDDP+VNB) arm had a higher incidence of haematological toxicity, except for grade 3 and 4 anaemia, which was higher for the CDDP+VNB (IFOS+EPI) arm Non-haematological toxicity: In step 1, the IFOS+EPI (CDDP+VNB) arm had a higher incidence of non-haematological toxicity, except for grade 3 and 4 diarrhoea, which was equivalent for both groups. In step 2, the IFOS+EPI (CDDP+VNB) arm had a higher incidence of non-haematological toxicity, except for grade 3 and 4 nausea/vomiting, which was more evident with CDDP+VNB (IFOS+EPI)
Comella *et al.*, 1996[39] Design: Phase III, multicentre, randomised trial. ITT analysis Intervention: CDDP 40 mg/m^2 with VP-16 100 mg/m^2 (53 eligible patients) every 28 days; CBDCA 250 mg/m^2 with CDDP 30 mg/m^2, VP-16 100 mg/m^2 and VNB 30 mg/m^2 (52 eligible patients) every 28 days Patients: Stage IIIB or IV NSCLC (n = 105 eligible patients) Jadad quality score: 3/5	Complete response: CDDP+VP-16, 0% CBDCA+CDDP+VNB, 2% Partial response: CDDP+VP-16, 28% CBDCA+CDDP+VNB, 23% Overall response rate: CDDP+VP-16 , 28% (95% CI, 17% to 42%) CBDCA+CDDP+VNB, 25% (95% CI, 13% to 37%) (Mantel-Haenszel, p = 0.7)	Median survival time: CDDP+VP-16, 31 weeks CBDCA+CDDP+VNB, 27 weeks (p = NS)	QoL: No QoL data	Haematological toxicity: Toxicity was most evident in the CBDCA+CDDP+VNB arm, except for grade 3 and 4 anaemia, which was equally evident in both groups. Neutropenia (p < 0.001) and thrombocytopenia (p < 0.001) were significantly more frequent with CBDCA+CDDP+VNB Non-haematological toxicity: Grade 4 non-haematological toxicity did not occur, except for nausea/vomiting, which was equally evident in both arms. Grade 3 non-haematological toxicity was more evident with CBDCA+CDDP+VNB, except for ototoxicity, which was more evident with CBDCA+CDDP+VNB

continued

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Crawford *et al.*, 1996[40] Design: Phase III, multicentre, randomised trial. ITT analysis Intervention: VNB 30 mg/m^2 (143 patients) every week, 5FU 425 mg/m^2 with LV 20 mg/m^2 (68 patients) every 4 weeks Patients: Stage IV NSCLC (n = 216) Jadad quality score: 3/5	Complete response: VNB, 0% 5FU+LV, 0% Partial response: Not reported Overall response rate: VNB, 12% 5FU+LV, 3% (p = NS)	Median survival time (estimated): VNB, 30 weeks 5FU+LV, 22 weeks (log-rank test, p = 0.03) One-year survival rate: VNB, 25% 5FU+LV, 16% (Cox model, p = 0.06)	QoL: No data provided. Minimal changes in scores were observed overall for each dimension for the two treatments. A trend toward improved distress scores was evident for VNB patients compared with 5FU+LV patients. Physical functioning was better for 5FU+LV patients than VNB patients. There was some improvement in global QoL scores for both arms. No treatment differences were noted in relation to relief of cancer symptoms. During the first 8 weeks of treatment, 36% of VNB patients and 45% of 5FU+LV patients showed an improvement in all baseline cancer symptoms	Haematological toxicity: Grade 3 and 4 granulocytopenia and granulocytopenic infections were much more evident in the VNB arm, whereas grade 3 and 4 thrombocytopenia and anaemia were marginally more evident in the 5FU+LV arm Non-haematological toxicity: No grade 4 toxicities for the VNB arm were reported. All grade 3–4 non-haematological toxicities were more evident in the 5FU+LV arm
Depierre *et al.*, 1994[41] Design: Phase III, multicentre, randomised trial. ITT analysis Intervention: VNB 30 mg/m^2 (119 patients) weekly, VNB 30 mg/m^2 with CDDP 80 mg/m^2 (121 patients) every 3 weeks Patients: Stage IIIA/B or IV NSCLC (n = 240) Jadad quality score: 3/5	Complete response: VNB, 1% VNB+CDDP, 2% Partial response: VNB, 2% VNB+CDDP, 41% Overall response rate: VNB, 16% (95% CI, 9% to 22%) VNB+CDDP, 43% (95% CI, 34% to 52%) (p = 0.0001)	Median survival time: VNB, 32 weeks VNB+CDDP, 33 weeks (log-rank test, p = 0.48) One-year survival: Not reported	QoL: No QoL data	Haematological toxicity: WHO grades for haematological toxicities were not reported. Haematological toxicity was much more evident in the VNB+CDDP arm Non-haematological toxicity: Toxicity was much more evident in the VNB+CDDP arm, except for grade 3–4 constipation, which was equally evident in both arms

continued

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Furuse *et al.*, 1996[42] Design: Phase III, crossover, multicentre, randomised trial. ITT not specified Intervention: • VNB arm: VNB 25 mg/m^2 weekly (103 patients); patients not responding after 4 cycles were switched to VDS 3 mg/m^2 with CDDP 80 mg/m^2, every 4 weeks for last 2 cycles • VDS arm: VDS 3 mg/m^2 weekly (101 patients); patients not responding after 4 cycles were switched to VNB 20 mg/m^2 with CDDP 80 mg/m^2, every 4 weeks for last 2 cycles Patients: Stage IIIB or IV NSCLC (n = 204) Jadad quality score: 3/5	Complete response: VNB arm, 0% VDS arm, 0% Partial response: VNB arm, 32 (31.1%; 95% CI, 22.3% to 40.8%) VDS arm, 9 (8.9%; 95% CI, 4.2% to 16.2%) Overall response: VNB arm, 13/49 (26.5%; 95% CI, 15% to 40.7%) VDS arm, 0/33 (0%)	Median survival time: VNB arm, 52.4 weeks VDS arm, 43.6 weeks (p = 0.3839) One-year survival: Not reported	QoL: No QoL data	Haematological toxicity: • For the monotherapy part of the treatment, haematological toxicity was much more evident in the VNB arm; this was significant only for grade 3–4 anaemia (p = 0.014) • With combination therapy, haematological toxicity was not significantly different in either arm, with leukocytopenia more evident in the VDS+CDDP arm and anaemia more evident in the VNB+CDDP arm Non-haematological toxicity: • For the monotherapy part of the treatment, non-haematological toxicity was more evident in the VDS arm, except for nausea/vomiting and local cutaneous reaction. Differences were significant only for alopecia (p = 0.001), peripheral neurotoxicity (p = 0.002) and local cutaneous reaction (p = 0.012) • With combination therapy, non-haematological toxicity was not significantly different, except for alopecia (VNB+CDDP, 0%; VDS+CDDP, 3%; p = 0.006)
Le Chevalier *et al.*, 1994[43] Design: Phase III, international, multicentre, randomised trial. ITT not specified Intervention: VNB 30 mg/m^2 with CDDP 120 mg/m^2 (206 patients) every 6 weeks, VDS 3 mg/m^2 with CDDP 120 mg/m^2 (200 patients) every 6 weeks, VNB 30 mg/m^2 (206 patients) weekly Patients: Stage III or IV NSCLC (n = 612) Jadad quality score: 3/5	Complete response: Not reported Partial response: Not reported Overall response: VNB+CDDP, 57 (30%) VDS+CDDP, 35 (19%) VNB, 28 (14%) VNB+CDDP had a significantly higher response rate than VDS+CDDP (χ^2 test, p = 0.02) and VNB (χ^2 test, p < 0.001)	Median survival time: VNB+CDDP, 40 weeks VDS+CDDP, 32 weeks VNB, 31 weeks • Unadjusted log-rank test: VNB+CDDP vs VDS+CDDP, p = 0.085 VNB+CDDP vs VNB, p = 0.045 • Centre-adjusted log-rank test: VNB+CDDP vs VDS+CDDP, p = 0.04 VNB+CDDP vs VNB, p = 0.01 One-year survival: Not reported	QoL: No QoL data	Haematological and non-haematological toxicities were more evident in the CDDP arms. Significance levels were not given

continued

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Lorusso *et al.*, 1995[44] Design: Phase III, multicentre, randomised trial. Only evaluable patients were analysed Intervention: VNB 25 mg/m^2 (35 patients) every week, VNB 25 mg/m^2 with CDDP 80 mg/m^2 (34 patients) every 3–4 weeks Patients: Inoperable NSCLC (n = 69) Jadad quality score: 2/5	Complete response: VNB, 0% VNB+CDDP, 0% Partial response: VNB, 4 (13%) VNB+CDDP, 13 (42%) (Fisher's exact test, p = 0.038) Overall response: not reported	Median survival time: VNB, 30 weeks VNB+CDDP, 38 weeks (p = NS) One-year survival: Not reported	QoL: No QoL data	Haematological toxicity: Toxicity was more evident in the VNB+CDDP arm, except for grade 3–4 leucopenia and grade 3–4 phlebitis, which were more evident in the VNB arm. Differences in toxicities were significant only for grade 3–4 anaemia (p < 0.01)
Martoni *et al.*, 1998[45] Design: Phase III, multicentre, randomised trial. ITT analysis Intervention: EPI 120 mg/m^2 with CDDP 60 mg/m^2 (102 eligible patients) every 21 days, VNB 25 mg/m^2 with CDDP 60 mg/m^2 (110 eligible patients) every 21 days Patients: Locally advanced or metastatic NSCLC (n = 212 eligible patients) Jadad quality score: 2/5	Complete response: EPI+CDDP, 1% VNB+CDDP, 2% Partial response: EPI+CDDP, 31% VNB+CDDP, 25% Overall response rate: EPI+CDDP, 32% (95% CI, 23% to 41%) VNB+CDDP, 27% (95% CI, 18% to 34%) (p = 0.45)	Median survival time: EPI+CDDP, 10.5 months (95% CI, 9.4 to 11.5 months) VNB+CDDP, 9.6 months (95% CI, 8.4 to 10.8 months) One-year survival rate: EPI+CDDP, 42% VNB+CDDP, 39% (p = NS) Two-year survival rate: EPI+CDDP, 15% VNB+CDDP, 8% (p = NS)	QoL: QoL was monitored via PS and symptoms Percentage of patients with improvement ≥ 10% in the Karnofsky PS: EPI+CDDP, 37% VNB+CDDP, 39% At least one symptom suffered before treatment: EPI+CDDP, 64% VNB+CDDP, 70% An improvement in at least one symptom, without worsening or the appearance of another: EPI+CDDP, 57% VNB+CDDP, 61% Differences were not significant	Haematological toxicity: Grade 3–4 leucopenia (p = 0.01) and thrombocytopenia (p = 0.02) were significantly more frequent in the EPI+CDDP arm. Grade 3–4 anaemia was more frequent in the VNB + CDDP arm (p = NS). Febrile neutropenia was more evident in the EPI+CDDP arm Non-haematological toxicity: The only grade 3 adverse events evident were nausea/vomiting, which were more evident in the EPI+CDDP arm (p = NS). Grade 1–2 adverse events were more evident for the EPI+CDDP arm, except for vomiting/nausea and renal toxicity. Differences were not significant, except for Grade 1–2 alopecia, which was significantly more evident in the EPI+CDDP arm (p = 0.001)

continued

Patient survival

The five RCTs comparing VNB with different doses and combinations of VNB and other interventions reported higher but non-significant median survival times for patients on VNB+CDDP compared with those on VNB alone.[36,37,41,43,44] Similarly, higher but non-significant median survival times were shown for VNB+CDDP+MITO and CDDP+IFOS+VNB compared with CBDCA+VNB.[36,37] The two RCTs reporting studies that operated some form of crossover design, whereby patients changed interventions following prespecified criteria,[38,42] reported higher non-significant median survival times for the VNB arms. Of the remaining six RCTs,[39,40,45–48] three showed that the VNB combination arm had higher median survival times,[40,46,47] with two significantly higher by 2 months in both

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
Pérol *et al.*, 1996[46] Design: Phase II, open, multicentre, randomised trial. ITT used for survival and toxicity only Intervention: CDDP 120 mg/m^2 with MITO 8 mg/m^2 and VDS 3 mg/m^2 on days 1, 29 and 71 for 5 weeks and then every 2 weeks up to 15 weeks (113 patients); CDDP 120 mg/m^2 with MITO 8 mg/m^2 and VNB 25 mg/m^2 on days 1, 29 and 71 for 16 weeks (114 patients) Patients: Stage III or IV NSCLC (n = 227) Jadad quality score: 2/5	Complete response: CDDP+MITO+VDS, 2% CDDP+MITO+VNB, 4% Partial response: CDDP+MITO+VDS, 15% CDDP+MITO+VNB, 21% Overall response rate: CDDP+MITO+VDS, 17% (95% CI, 10% to 24%) CDDP+MITO+VNB, 25% (95% CI, 17% to 32%) (p = 0.15)	Median survival time: CDDP +MITO+VDS, 33.4 weeks CDDP+MITO+VNB, 34.5 weeks (log-rank test, p = 0.2) Overall 2-year survival rate: CDDP+MITO+VDS, 15.6% CDDP+MITO+VNB, 9% (p = 0.13)	QoL: No QoL data	Haematological toxicity: Grade 3–4 anaemia (p < 0.01) and neutropenia (p < 0.01) were significantly more evident in the CDDP+MITO+VNB arm. Grade 3–4 thrombocytopenia was more evident in the CDDP+MITO+VDS arm, but the difference was not significant Non-haematological toxicity: Grade 2–4 sepsis (p < 0.03) and local reaction (p < 0.05) were more evident in the CDDP+MITO+VNB arm. Grade 2–4 neurological events (p < 0.01) and nausea/ vomiting (p = NS) were more evident in the CDDP +MITO+VDS arm. Grade 3–4 renal and cardiac events were more evident, though not significantly, in the CDDP+MITO+VDS arm
Wozniak *et al.*, 1998[47] Design: Phase III, multicentre, randomised trial. ITT not used for response rates Intervention: VNB 25 mg/m^2 weekly with CDDP 100 mg/m^2 every 4 weeks (206 patients), CDDP 100 mg/m^2 every 4 weeks (209 patients) Patients: Stage IIIB or IV NSCLC (n = 415) Jadad quality score: 2/5	Complete response: VNB+CDDP, 2% CDDP, 0% Partial response: VNB+CDDP, 24% CDDP, 12% Overall response rate: VNB+CDDP, 26% CDDP, 12% (p < 0.001)	Median survival time: VNB+CDDP, 8 months CDDP, 6 months (p = 0.0018) One-year survival rate: VNB+CDDP, 36% CDDP, 20% Two-year survival rate: VNB+CDDP, 12% CDDP, 6%	QoL: No QoL data	Haematological toxicity: Toxicity was more evident in the VNB+CDDP arm Non-haematological toxicity: Toxicity was more evident with VNB+CDDP, except for Grade 3–4 nausea/ vomiting and renal events, which were equally evident in both arms

continued

cases.[40,47] In two RCTs, the VNB+CDDP and VNB+CDDP+CBDCA+VP-16 combinations showed non-significant decreases in median survival time.[39,45] The comparison of VNB with BSC showed longer median survival time for patients receiving VNB.[48]

Six RCTs assessed patient survival to 1 or 2 years.[36,40,45–48] Of the five RCTs reporting patient survival at 12 months, two showed higher proportions of patients surviving in VNB (25%) and VNB+CDDP (36%) arms when compared with F5U+LV (16%) and CDDP alone (20%), respectively.[40,47] When compared with CDDP+MITO+VDS and EPI+CDDP, lower proportions of patients receiving CDDP+IFOS+VNB, CBDCA+VNB or VNB+CDDP survived to 12 months.[36,45] The RCT comparing VNB with BSC reported

TABLE 6 contd *Summary of evidence of the effectiveness of vinorelbine in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
ELVIS Group, 1999[48] Design: Phase III, multicentre, randomised trial. Non-ITT Intervention: VNB (intravenous) 30 mg/m^2 on days 1 and 8 of 21-day cycle (76 patients), BSC (78 patients) Patients: Stage IIIB or IV (n = 154) Jadad quality score: 3/3	Complete response: VNB, 1/76 (1%) Partial response: VNB, 14/76 (18%) Objective response rate: VNB, 19.7% (95% CI, 11.5% to 30.5%)	Median survival time: VNB, 28 weeks BSC, 21 weeks Six-month survival rate: VNB, 55% BSC, 41% One-year survival rate: VNB, 32% BSC, 14%	QoL: • At baseline, there was no difference between groups, although suboptimal values were evident for global health status: VNB, 36.8% BSC, 31.5% • After 6 cycles, EORTC functional scales were better for the VNB arm, significantly for cognitive function (p = 0.02). Symptom scales were significantly better for the VNB arm on pain (p = 0.02), but significantly worse on constipation (p = 0.002) and nausea and vomiting (p = 0.07). For LC13 scales, the VNB arm was significantly better for dyspnoea (p = 0.05) and pain medication (p = 0.01), but worse for peripheral neuropathy (p = 0.04) and hair loss (p = 0.0001)	Treatment was stopped in 5 patients due to severe toxic events. WHO grade 3 constipation occurred in 3 patients, grade 4 constipation in 1 patient, and grade 2 heart toxicity in 1 patient. Other relevant toxic events, not including treatment stoppage, were grade 4 leucopenia (1 patient), grade 4 neutropenia (3 patients), grade 3 vomiting (1 patient) and grade 3 alopecia (3 patients)

PS, performance status; ELVIS, Elderly Lung Cancer Vinorelbine Italian Study

higher 6- and 12-month survival in the VNB arm.[48] Over 24 months, the proportion of patients surviving was higher for patients receiving VNB+CDDP compared with CDDP,[47] but lower for patients receiving VNB+CDDP compared with EPI+CDDP.[45] In only two RCTs was the statistical significance of any differences tested, with non-significant differences shown.[40,45]

Quality of life

Of the 13 RCTs, only three assessed the effect of interventions on QoL.[40,45,48] QoL was assessed using South West Oncology Group (SWOG) questionnaires and physician assessments,[40] by evaluating performance status and symptoms,[45] or using the EORTC QLQ-C30 and LC13 scales.[48] Two of the three RCTs showed limited difference in QoL between VNB and 5FU+LV,[40] and between VNB+CDDP and EPI+CDDP,[45] but did not use the best current measure and gave little detail of analysis, making evaluation difficult. One RCT reported that fewer QoL questionnaires were completed by the 5FU group due to dropout from the study (survival was poorer).

The comparison of VNB with BSC did use the EORTC QLQ-C30 and LC13 measures, and showed significant improvements for VNB on cognitive function (p = 0.02), pain (p = 0.02), pain and medication (p = 0.01) and dyspnoea (p = 0.05), but significant worsening on constipation (p = 0.002), nausea and vomiting (p = 0.07), peripheral neuropathy (p = 0.04) and hair loss (p < 0.001).

Adverse effects

Adverse events, haematological and non-haematological, were reported by all 13 RCTs. In the five RCTs comparing two VNB combination arms,[36,37,41,43,44] three found higher reporting of

haematological and non-haematological adverse events in combinations including VNB and CDDP.[36,41,44] The other RCTs comparing VNB combinations found limited difference in adverse effects between interventions,[37,43] although higher rates of haematological effects were reported for CBDCA+VNB when compared with CDDP+MITO+VNB.[37] In the two RCTs adopting forms of crossover design, the RCT comparing CDDP+VNB followed by IFOS+EPI against IFOS+EPI followed by CDDP+VNB found the latter combination to have higher rates of haematological and non-haematological adverse effects following each phase of treatment.[38] The RCT comparing VNB (followed by VDS+CDDP) with VDS (followed by VNB+CDDP) found higher reporting of haematological events following VNB and higher reporting of non-haematological events after VDS during the monotherapy comparison.[42] When assessed following the additional combination therapy, there were no differences in adverse events. In the three RCTs comparing combinations of VNB and CDDP,[45–47] patients receiving EPI+CDDP suffered significant haematological toxicity from leucopenia and thrombocytopenia as well as non-haematological toxicities in the form of alopecia, compared with those receiving VNB+CDDP.[45] Patients receiving CDDP+MITO+VNB suffered significantly more anaemia, neutropenia, sepsis and local reaction, compared with those receiving CDDP+MITO+VDS.[46] In the RCT comparing VNB+CDDP and VNS+CDDP, haematological adverse effects were more frequent in the former group and non-haematological adverse effects were more frequent in the latter.[47] When compared with the 5FU+LV group, patients receiving VNB suffered more granulocytopenia and related infection, while patients receiving 5FU+LV suffered higher rates of thrombocytopenia and anaemia.[40] Patients receiving 5FU+LV were more likely to suffer non-haematological adverse effects. In the trial comparing CDDP+VP-16 and CBDCA+CDDP+VP-16+VNB, the latter group suffered a higher incidence of haematological and non-haematological adverse effects.[39] In the RCT comparing VNB with BSC, five patients stopped treatment due to severe toxic events. Adverse effects included constipation, heart toxicity, leucopenia, neutropenia, vomiting and alopecia.

Summary of the use of vinorelbine in NSCLC

Survival

Compared with BSC, VNB improved survival by about 7 weeks; 1-year survival was 32% in patients on VNB, compared with 14% in those on BSC.

Comparisons with other drugs gave a mixed picture. VNB in combination with CDDP was better than CDDP alone, producing a 2-month gain in average survival, a 1-year survival rate of 36% compared with 20%, and 2-year survival of 12% versus 6%, respectively. Survival on VNB was also better than with the 5FU combination, but 5FU is a drug not normally used in patients with lung cancer. In most of the other trials, there were improvements with VNB combinations, but these were usually not statistically significant.

Quality of life

Disappointingly, only three of the VNB trials included assessment of QoL, and only one used the current best method. That trial showed that, compared with BSC, QoL was better for most but not all indicators in patients receiving VNB. Hence, the gain in survival was accompanied by an improved QoL.

Combined therapies

Quantity and quality of research for combined therapies in NSCLC

Five RCTs met the inclusion criteria for the review and are shown in *Table 7* and appendix 9.[49–53]

Study quality

Four of the five published RCTs had Jadad quality scores of 2 out of 5,[49,51–53] while the other RCT had a Jadad quality score of 3 out of 5.[50] All the studies were judged to be of reasonably good quality. None of the studies discussed blinding; however, blinding would have been difficult given the comparators used. The RCT scoring 3 on the Jadad scale adequately described the method of randomisation used as well as withdrawals and dropouts.[50] The RCTs with Jadad scores of 2 provided descriptions of the withdrawals and dropouts, but there was limited discussion of the method of randomisation.[49,51–53] In addition, the poor-quality RCTs lacked information on patient characteristics,[52] reporting of statistical analysis[52] or definition of outcomes.[51]

The main primary end-points used in the studies were survival and tumour response rates, with secondary end-points being time to event parameters, toxicity, QoL and survival.

Assessment of clinical effectiveness of combination therapies in NSCLC

There were five different variations of the combination therapy RCTs. Two RCTs compared CDDP+GEM+VNB with either CDDP+EPI+

TABLE 7 *Summary of evidence of the effectiveness of combined therapies in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	Survival	Other outcome	
Comella *et al.*, 1999[49] Design: Phase II, randomised trial. ITT Intervention: CDDP 50 mg/m^2 with GEM 100 mg/m^2 and VNB 25 mg/m^2 (57 patients), CDDP 80 mg/m^2 with EPI 80 mg/m^2, VDS 3 mg/m^2 and LON 150 mg/m^2 (54 patients) Patients: Stage IIIB or IV NSCLC (n = 111) Jadad quality score: 2/5	Complete response: CDDP+GEM+VNB, 2% CDDP+EPI+VDS+LON, 4% Partial response: CDDP+GEM+VNB, 58% CDDP+EPI+VDS+LON, 33% Overall response rate: CDDP+GEM+VNB, 60% (95% CI, 46% to 72%) CDDP+EPI+VDS+LON, 37% (95% CI, 24% to 51%)	Median survival time: CDDP+GEM+VNB, 50 weeks (95% CI, 41 to 58 weeks) CDDP+EPI+VDS+LON, 33 weeks (95% CI, 24 to 41 weeks) One-year survival rate: CDDP+GEM+VNB, 48% CDDP+EPI+VDS+LON, 29% Two-year survival rate: CDDP+GEM+VNB, 19% CDDP+EPI+VDS+LON, 0%	Improved QoL score: CDDP+GEM+VNB, 59% CDDP+EPI+VDS+LON, 39%	Haematological toxicity: Grade 3–4 toxicity was more evident in the CDDP+GEM+VNB arm, except for anaemia, which was more evident in the CDDP+EPI+VDS+LON arm. No significance levels were given Non-haematological toxicity: Grade 3–4 toxicity was more evident in the CDDP+EPI+VDS+LON arm, except for renal, neuropathy, mucositis and diarrhoea toxicities, which were not evident in either arm, as well as fatigue, which was equally evident in both arms
Comella *et al.*, 2000[50] Design: Phase III, randomised trial. Interim analysis. ITT Intervention: CDDP 50 mg/m^2 with GEM 1000 mg/m^2 and VNB 25 mg/m^2 (60 patients), CDDP 100 mg/m^2 with GEM 1000 mg/m^2 (60 patients), CDDP 120 mg/m^2 with VNB 30 mg/m^2 (60 patients) Patients: Stage IIIB or IV NSCLC (n = 180) Jadad quality score: 3/5	Complete response: CDDP+GEM+VNB, 3% CDDP+GEM, 0% CDDP+VNB, 0% Partial response: CDDP+GEM+VNB, 43% CDDP+GEM, 30% CDDP+VNB, 25% Overall response rate: CDDP+GEM+VNB, 47% (95% CI, 34% to 60%) CDDP+GEM, 30% (95% CI, 19% to 43%) CDDP+VNB, 25% (95% CI, 15% to 38%)	Median survival time: CDDP+GEM+VNB, 51 weeks CDDP+GEM, 42 weeks CDDP+VNB, 35 weeks One-year survival rate: CDDP+GEM+VNB, 45% CDDP+GEM, 40% CDDP+VNB, 34%	QoL: No QoL data presented	Haematological toxicity: Grade 3–4 toxicity was more evident with CDDP+VNB, except for thrombocytopenia, which was more evident in the CDDP+GEM arm Non-haematological toxicity: Grade 3–4 toxicity was more evident in the CDDP+VNB arm, except for mucositis and diarrhoea, which were more evident in the CDDP+GEM arm. The CDDP+GEM+VNB arm had the least incidence of non-haematological toxicity. Grade 3–4 neutropenia and vomiting were significantly greater in the CDDP+VNB arm

continued

TABLE 7 contd *Summary of evidence of the effectiveness of combined therapies in treating NSCLC in published RCTs*

Author, year and study details	Key outcome measures			Adverse effects
	Tumour response	**Survival**	**Other outcome**	
Kosmidis *et al.*, 2000[51] Design: Phase III, randomised trial. Preliminary results. ITT not specified Intervention: PAX 200 mg/m^2 with CBDCA (63 patients), PAX 200 mg/m^2 with GEM 1000 mg/m^2 (64 patients) Patients: Stage IIIA (inoperable), stage IIIB or IV NSCLC (*n* = 127) Jadad quality score: 2/5	Complete response: PAX+CBDCA, 0% (95% CI, 0% to 7.7%) PAX+GEM, 4.2% (95% CI, 0.5% to 14.3%) Partial response: PAX+CBDCA, 21.8% (95% CI, 9.4% to 33.9%) PAX+GEM, 33.3% (95% CI, 20.4% to 48.4%)	Median survival: Not reported One-year survival: Not reported	QoL: Not reported	Haematological toxicity: Toxicity was more evident in the PAX+CBDCA arm, although the difference was not statistically significant Non-haematological toxicity: Toxicity was more evident in the PAX+CBDCA arm, except for myalgia, arrhythmia and cutaneous events, which were greater in the PAX+GEM arm. No significance levels were reported
Perry *et al.*, 2000[52] Design: Phase II, randomised trial. ITT not specified Intervention: PAX 250 mg/m^2 with IFOS 1.6 g/m^2 (48 patients), VNB 30 mg/m^2 with IFOS 1.6 g/m^2 (45 patients) Patients: Stage IIIB or IV NSCLC (*n* = 93) Jadad quality score: 2/5	Complete response: PAX+IFOS, 4% VNB+IFOS, 4% Partial response: PAX+IFOS, 33% VNB+IFOS, 24% Overall response rate: PAX+IFOS, 38% (95% CI, 24% to 53%) VNB+IFOS, 31% (95% CI, 18% to 47%)	Median survival time: PAX+IFOS, 8.5 months VNB+IFOS, 7.4 months (95% CI, 5.3 to 13.3 months) One-year survival rate (estimated): PAX+IFOS, 35% (95% CI, 24% to 52%) VNB+IFOS, 38% (95% CI, 26% to 55%)	QoL: No data collected	Haematological toxicity: Grade 3–4 toxicity was more evident in the VNB+IFOS arm, except for hyperglycaemia, which was more evident in the PAX+IFOS arm Non-haematological toxicity: Toxicity was more evident in the VNB+IFOS arm, except for dyspnoea, which was more evident in the PAX+IFOS arm. No significance levels were provided
Frasci *et al.*, 2000[53] Design: Phase III, randomised, trial. Interim analysis (study not powered to detect significant differences). ITT analysis specified Intervention: GEM 1200 mg/m^2 with VNB 30 mg/m^2 (60 patients), VNB 30 mg/m^2 (60 patients) Patients: Stage IIIB or IV NSCLC (*n* = 120) Jadad quality score: 2/5	Complete response: GEM+VNB, 0% VNB, 0% Partial response: GEM+VNB, 22% VNB, 15% Overall response rate: GEM+VNB, 22% (95% CI, 12% to 34%) VNB, 15% (95% CI, 7% to 27%)	Median survival time: GEM+VNB, 29 weeks VNB, 18 weeks Six-month survival rate (estimated): GEM+VNB, 56% VNB, 32% One-year survival rate (estimated): GEM+VNB, 30% VNB, 13%	QoL: Almost 60% of the GEM+VNB patients did not show impairment of the QoL score during treatment, compared with approximately 40% of the VNB arm. There was insufficient reporting of QoL measures, and no significance levels were provided	Haematological toxicity: Grade 3–4 toxicity was more evident in the GEM+VNB arm Non-haematological toxicity: Grade 3–4 toxicity was more evident in the GEM+VNB arm. No significance levels were provided

VDS+LON[49] or CDDP+GEM and CDDP+VNB.[50] One RCT compared PAX+IFOS with VNB+IFOS,[52] another GEM+VNB with VNB,[53] while the fifth compared PAX+CBDCA with PAX+GEM.[51]

Patient survival

The two RCTs comparing CDDP+GEM+VNB with other combination therapies reported greater median and 1-year survival for the CDDP+GEM+VNB arm as opposed to the comparator arm,[49,50] significantly for median survival when compared with CDDP+EPI+VDS+LON.[49] One RCT comparing CDDP+ GEM+VNB with CDDP+GEM and CDDP+VNB showed that, after CDDP+GEM+VNB, the GEM+CDDP arm had the highest survival rate.[50] The trial that compared GEM+VNB with VNB described a longer median survival time for the combination arm.[53] Of the two remaining RCTs, one did not report survival rates;[51] while the other RCT reported a higher survival for the PAX+IFOS arm compared with the VNB+IFOS arm, the difference was not significant.[52]

Quality of life

Four RCTs said that QoL was an outcome measure,[49,50,52,53] although only two RCTs actually reported QoL data.[49,53] One RCT found that there was a greater improvement in QoL scores for the CDDP+VNB+GEM arm than CDDP+EPI+VDS.[49] The other trial reported that approximately 60% of patients in the GEM+VNB arm did not show impairment, compared with 40% in the VNB arm, although QoL reporting was suggested to be insufficent.[53]

Adverse effects

Haematological and non-haematological adverse events were reported in all five trials, although the significance of any differences was not stated. In the RCT comparing CDDP+GEM+VNB with CDDP+EPI+VDS, haematological toxicities were generally more common in the former, while non-haematological toxicities were more common in the latter.[49] When compared with CDDP+GEM+VNB and CDDP+VNB, the CDDP+GEM arm was reported to have more haematological and non-haematological adverse effects.[50] CDDP+VNB had the lowest incidence of haematological adverse effects, while CDDP+ GEM+VNB had the lowest incidence of non-haematological adverse effects. Comparison of adverse effects resulting from PAX+CBDCA with those from PAX+GEM showed that the former had higher haematological and non-haematological toxicity than the latter.[51] Similarly, VNB+IFOS had higher incidences of haematological and non-haematological adverse effects than PAX+IFOS.[52] The GEM+VNB arm had higher incidences of haematological and non-haematological reported events than the VNB arm alone.[53]

Summary of use of combination therapies in NSCLC

Survival

The combination of CDDP+VNB+GEM gave longer survival by 11–19 weeks than older combinations, with 19% more patients surviving to 1 year (48% in CDDP+VNB+GEM arm vs 29% in CDDP+EPI+VDS+LON arm). Triple therapy with CDDP combined with both VNB and GEM, versus CDDP+GEM, and versus CDDP+VNB, gave survival times of 51 weeks, 42 weeks and 35 weeks, respectively.[50]

Adding GEM to VNB produced a survival gain of 11 weeks (29 vs 18 weeks). Survival data vary among trials due to patient mix.

Quality of life

Only two trials reported on QoL: one trial used a non-standard scale, and the other gave sparse detail. The information is insufficient to make any safe conclusions.

Other issues

Histology

NSCLC consists mainly of three subtypes: squamous cell, adenocarcinoma and large cell. Almost all the RCTs gave the breakdown, but only three of all the studies gave results by subtype. Response rates were not different in two trials (Comella and co-workers,[49] and Crawford and colleagues[40]), and were better in squamous cell cancer in the third (Giaccone and co-workers[35]).

However, one curious finding was that squamous cell cancers were present in about 34% of patients in the GEM and PAX studies, but in about 53% in the VNB trials.

Problems with the evidence

There were four main problems with the evidence. Firstly, there was a great profusion of regimens and comparators, making it difficult to compare the new drugs. Head-to-head comparisons would have helped.

Secondly, there were few good QoL studies, and these were mostly for GEM.

Thirdly, the trials had few patients over the age of 75 years, whereas in the UK almost half of patients with lung cancer are over that age.

Lastly, trials may not reflect real life, and this will affect costs derived from them (discussed in chapter 4). In practice, patients may get 'N of 1' trials, wherein chemotherapy is continued only if the patient shows symptomatic and radiological response, whereas in protocol-driven trials, chemotherapy regimens will be decided in advance and given to the limits of tolerance. In routine practice, patients who do not respond may get only one or two courses, and will therefore have fewer side-effects. This makes chemotherapy more efficient.

Chapter 4

Economic analysis

Literature review

A literature review was conducted to identify economic studies or costing papers on the use of PAX, DOC, GEM and VNB in NSCLC. None were found for these drugs in SCLC. Twenty studies were found for inoperable stage III and stage IV NSCLC. Four of these were reviews, none of which led to the identification of published studies that had not previously been identified. Hence, the reviews were bypassed in favour of appraising the underlying studies.

Seven studies considered GEM, five considered VNB, one included both GEM and VNB regimens, and two considered PAX only. The most recent study considered GEM, VNB and PAX.[54] None of the economic or costing studies considered DOC.

The characteristics of the 16 economic or costing studies are presented in appendix 10.

Summary of findings of cost-effectiveness

Appendix 11 describes the cost-effectiveness results of the economics studies. Conclusions from these studies are summarised below. Seven first-line regimens are considered, including the three single agents and GEM, VNB and PAX together with CDDP, the most commonly used drugs in combinations.

Single-agent GEM

Berthelot and co-workers (2000)[54] determined that GEM was dominated by (i.e. was more expensive and less effective than) VNB+CDDP and had an incremental cost per life-year saved (LYS) of Can$17,400, compared with VNB. GEM was cost-saving compared with PAX+CDDP (at three PAX doses) and VDS+CDDP. At low doses (1000–1250 mg/m^2 per administration), GEM has also been shown to lead to only a small incremental cost over the Canadian BSC practice.[55] Berthelot and co-workers[54] reported a incremental cost per LYS of Can$6800 compared with BSC. Assuming equal efficacy and excluding drug costs, single-agent GEM was reported to lead to cost savings compared with VP-16+CDDP and IFOS+VP-16.[56,57] This difference is largely due to the use of GEM on an outpatient basis, whereas the other regimens are given as inpatient therapy.[58]

GEM with CDDP

Marginal analysis, in terms of additional cost to obtain a response, favoured GEM+CDDP over MITO+IFOS+CDDP, VP-16+CDDP and VNB+CDDP.[59] However, these calculations were based on the cost per response, assuming a higher response rate for GEM+CDDP. To estimate differences in cost-effectiveness, we need more information than just response, such as good evidence that response was highly correlated with patient-based outcomes like survival. Another study found there to be no statistical difference in direct costs between GEM+CDDP and VP-16+CDDP, assuming equal efficacy.[60]

Palmer and Brandt (1996)[59] included a comparison of GEM+CDDP and VNB+CDDP regimens. However, efficacy data and hence cost-effectiveness were again defined in terms of tumour response. GEM+CDDP (based on a response rate of 54%) was more cost-effective, with the incremental cost-effectiveness ratio (ICER) of VNB+CDDP (response rate of 35%) reported as 46.2 million lira per response. Direct costs across the two regimens were similar.

Single-agent VNB

Berthelot and co-workers[54] reported that VNB had an incremental cost of Can$1900, compared with BSC, and dominated VDS+CDDP and VP-16+CDDP. VNB had a lower cost than all the other regimens considered, bar VBL+CDDP. Evans (1996)[61] found single-agent VNB to be cost-saving compared with Canadian BSC, while delivering an average of 0.28 additional LYS per person. The cost saving per case was Can$1447.

VNB with CDDP

Berthelot and co-workers[54] found that VNB+CDDP dominated GEM, VP-16+CDDP and VDS+CDDP. It had an incremental cost of Can$8000 per LYS compared with single-agent VNB and Can$4100 compared with BSC. It was less expensive than PAX+CDDP (at three PAX doses).

Previous results reporting comparisons with VDS+CDDP have been mixed. While some authors have reported VNB+CDDP versus VDS+CDDP to be cost-additive, with an ICER of $15,500 per LYS,[62,63] others have calculated a cost saving with VNB+CDDP versus VDS+CDDP, with an ICER of Can$6000–7000 per LYS.[64,65] The last two papers are similar.

Smith and colleagues (1995)[62] and Hillner and Smith (1996)[63] reported VNB+CDDP as having an ICER of $17,700 per LYS compared with single-agent VNB. The authors' view was that the increase in efficacy afforded by VNB+CDDP over VNB is at reasonable cost. It should be noted that these are very similar papers by the same first two authors.

In further comparisons with a number of different regimens using BSC as the base case, single-agent VNB and VNB+CDDP (administered on an outpatient basis) were both cost-saving. Cost savings per case were Can$1447 for VNB and Can$473 for VNB+CDDP (outpatient). When VNB+CDDP was administered on an inpatient basis, the regimen incurred an incremental cost compared with BSC, VBL+CDDP and VP-16+CDDP. However, it also delivered the highest LYS (0.44 greater than BSC) and, when compared with BSC, had an ICER of Can$5551–6386 per LYS.[64,65]

When combined with MITO, VNB+CDDP produced a more favourable cost-per-response ratio than MITO+VDS+CDDP, based on a superior response rate.[66]

Single-agent PAX

A Canadian study that compared single-agent PAX with BSC found an incremental cost of Can$3375 and an ICER of Can$4778 per LYS (based on increased survival of 7.9 months with the PAX regimen) at baseline.[67]

PAX with CDDP

Berthelot and co-workers[54] found PAX+CDDP (at three PAX doses) to be the most expensive of the regimens considered but also the one that led to the longest survival gains. Thus, it was not dominated by any of the other regimens considered in their study. Incremental costs ranged from Can$15,400 to Can$27,000 (depending on dose) compared with BSC.

Annemans and colleagues (1999)[68] compared PAX+CDDP with VM-26+CDDP. There was an incremental cost for the PAX+CDDP regimen of US$2311 on average, but cost-effectiveness ratios were presented only in terms of cost per responder. The cost per responder was US$21,011 for PAX+CDDP, compared with US$27,266 for the VM-26+CDDP regimen (based on response rates of 37% and 26%, respectively). Again, response is not as satisfactory a measure of efficacy as survival or QoL.

Conclusion

The economic evaluations reported above have mostly been centre- or person-led, mainly by centres in the USA and Canada. No UK economic evaluations were identified. VNB is standard treatment in Canada and has been reported to deliver cost savings or low incremental cost compared with BSC. GEM and PAX have also led to small but acceptable incremental costs over BSC. However, in Phase III trials, there have been no direct comparisons of efficacy between VNB or GEM, the two most evaluated regimens; nor have there been comparisons with PAX or DOC. Studies evaluating DOC mostly see it as second-line therapy. The Canadian authors have acknowledged that side-by-side comparisons in trials are needed.

Validity

Internal validity

Internal validity appears to be reasonable in the economic studies and is summarised in appendix 12. There are two types of comparators in the studies: BSC or another alternate chemotherapy regimen. GEM is compared against BSC, VP-16+CDDP and IFOS+VP-16. GEM+CDDP is compared with VP-16+CDDP, VNB+CDDP and MITO+IFOS+CDDP. VNB is compared with BSC, VNB+CDDP, VDS+CDDP, VP-16+CDDP and VBL+CDDP. PAX is compared with BSC, and PAX+CDDP is compared with VM-26+CDDP. A range of comparisons are therefore made for GEM and VNB but restricted in terms of PAX.

Many of the studies did not include either costing of adverse events caused by chemotherapy or, in some cases, the cost of chemotherapy, and all the studies excluded non-health service costs. In the majority of studies, one-way sensitivity analysis was performed.

External validity (generalisability)

A number of European trials have included UK centres, and it seems reasonable to assume that efficacy results from these and other studies are applicable to patients in England and Wales. However, for cost-effectiveness purposes, two problems arise. Firstly, it is unreliable to convert

costs from elsewhere by a simple currency conversion. Secondly, treatments, including BSC, will follow US, Canadian or European practices, which may differ from those in the UK. Hence, the cost-effectiveness estimates from other countries may not apply to the UK. However, poorer survival rates for several cancers have been reported in the UK, compared with some other developed countries,[69] which suggests that there may be more scope for health gain here.

A more detailed checklist of the studies is represented in appendix 13. Neither the internal nor external validity have been evaluated in depth because of the lack of perceived usefulness served by such an exercise in which there are no published UK economic evaluations.

Estimating cost-effectiveness in the UK

In addition to the evidence from abroad, we have modelled the cost-effectiveness of the newer drug regimens in the UK. Meta-analysis was not possible for the reasons detailed in chapter 3, *Effectiveness*. Therefore, in order to overcome limitations in that data and to make the analysis more robust, three different modelling approaches have been used (see *Results of economic analysis* below).

Firstly, a series of pairwise comparisons of the drug regimens (or BSC) from actual published trials have been used to model cost-effectiveness (Model 1). In the majority of cases, the comparator was BSC.

Secondly, a cost-minimisation analysis was performed, assuming equal efficacy between regimens as a result of the lack of direct comparisons (Model 2).

Thirdly, a cost-effectiveness analysis (vs BSC) through synthesis of efficacy data by patient numbers was carried out (Model 3). Because no meta-analysis of efficacy was possible, this analysis is subject to a number of possible biases, discussed in detail on page 43.

Sources of data and costs used in the models

Costs

A number of impediments apply to the construction of a cost-effectiveness model for the UK. Firstly, there is a lack of readily available cost data. We have used data from a number of sources, including detailed work done by the Scottish Health Purchasing Information Centre (SHPIC) for an unpublished report. Costs for SHPIC were obtained by their costing unit based at Ninewells Hospital in Dundee, and involved 'bottom-up' costing. (Ninewells Hospital is classified in the Scottish Health Service Costs Book as a "large general major teaching hospital covering a full range of services (other than maternity in some cases) and with special units".) For BSC calculation, the SHPIC figure was based on a small cohort of patients. To supplement this, we obtained resource use data from a much larger cohort of patients from the Southeast Scotland Lung Study (SESLS) audit (unpublished). The latter data set did not provide cost data but only units of care, such as inpatient days and outpatient visits.

Cost structures may differ in a systematic way between Scotland, and England and Wales, and this risk is inherent in the model. However, it is likely that the increased NHS spend in Scotland goes toward more units of care, rather than increased cost per unit of care, and is likely to have less effect on cancer care than on, for example, rates for elective surgery. Nevertheless, differentials in bed numbers between Scotland and England, and among health authorities in England and Wales, need to be considered, because bed availability will affect what care can be given and may be a limiting factor in chemotherapy with present agents given on an inpatient basis.

All costs were obtained in or converted to 1999/2000 prices, using the published indices in Netten and Curtis (2000)[70] when necessary. Discounting was not deemed necessary in view of the short duration (< 1 year) of life-years gained.

Drug costs

All drug regimen costs were obtained from the British National Formulary (BNF) online in September 2000, using common doses extracted from the Phase III trials, and are shown in *Table 8*. Calculations were based on a body surface area of 1.7 m^2. If the published accounts of trials gave insufficient detail of dosages or number of courses, data were sought from the relevant industry submission or directly from the manufacturer.

The baseline scenarios in our modelling assume a certain degree of wastage through opened but unfinished vials for the different drugs. It was assumed the vials would not be reused once opened. Industry sources reported that it was common clinical practice to round down the dose if it was only slightly above vial size, to reduce wastage in terms of opened vials. We

TABLE 8 *Drug costs of chemotherapy regimens*

Drug	Regimen (mg/m^2)	Dose, by BSA (mg/m^2)	Constituents parts	Cost (£)
GEM (intravenous)	800	1360	One 1-g vial (£162.76) and two 200-mg vials (£32.55 each)	227.86
	1000	1700	One 1-g vial and four 200-mg vials	292.96
	1250	2125	Two 1-g vials and one 200-mg vial	358.07
VNB (intravenous)	25	42.5	10 mg/ml: one 5-ml vial (£147.06)	147.06
	30	51	10 mg/ml: one 5-ml vial and one 1-ml vial (£31.25)	178.31
PAX (intravenous)	135	229.5	6 mg/ml: two 16.7-ml vials (£374.00 each) and one 5-ml vial (£124.79)	872.79
	175	297.5	6 mg/ml: three 16.7-ml vials	1122.00
	200	340	6 mg/ml: three 16.7-ml vials and two 5-ml vials	1371.58
	250	425	6 mg/ml: four 16.7-ml vials and one 5-ml vial	1620.79
DOC (intravenous)	75	127.5	40 mg/ml: one 2-ml vial (£575.00) and three 0.5-ml vials (£175.00 each)	1100.00
	100	170	40 mg/ml: two 2-ml vials and one 0.5-ml vial	1325.00
CDDP (intravenous)	60	102	Two 50-mg vials (£17.00 each) and one 10-mg vial (£4.89)	38.89
	75	127.5	Two 50-mg vials and three 10-mg vials	48.67
	80	136	Three 50-mg vials	51.00
	100	170	Three 50-mg vials and two 10-mg vials	60.78
	120	204	Four 50-mg vials and one 10-mg vial	72.89

BSA, body surface area

Notes:

1. All drug costs were taken from the BNF, September 2000

2. All calculations were based on a BSA of 1.7 m^2

3. It is assumed that, once a vial is opened, it cannot be reused and shared among patients

have added a scenario to cover this in the sensitivity analysis. It is assumed a vial will be opened only if at least 20% (our assumption) of its contents are to be used on the patient. Below this amount, the vial will not be opened and the patient will be given a (marginally) smaller than recommended dose. Of course, this figure is arbitrary and will in practice depend on the opinion of the administering physician. In our scenario, the only regimens affected are the VNB-containing ones. Appendix 14 repeats *Table 8* with three additional columns indicating wastage (mg and %) and percentage of last vial used. It should also be noted that an average dose of, for example, 51 mg is just an average – many patients will need more or less. For the latter, there may be wastage from the 50-mg vial. To obtain more accurate costs would have required a survey of patients and doses, which was beyond the scope of this review.

Antiemetic and diuretic drugs used in published studies (dexamethasone, metoclopramide, cimetidine, methyl-prednisone and diphenhydramine; see appendix 15 for antiemetic regimens) all had negligible costs in the BNF and thus are excluded from the baseline analysis.

More modern drugs, such as ondansetron, are much more expensive but may also be more effective. A full review of the various antiemetics is outside the scope of this review, but work done for the SHPIC breast cancer report[71] showed that, although in terms of cost per drug, ondansetron was much more expensive – £6.75 for 4 mg for intravenous use, compared with £0.16 for metoclopramide – it was also much more effective. If successful treatment is defined as no vomiting during the 24 hours after chemotherapy, then success was achieved in 42% of patients given ondansetron and 25% of patients given metoclopramide. The cost per extra successfully treated patient was about £161. Clinical-based estimates could be obtained and incorporated. (A scenario is included in the sensitivity analysis of each model.)

In the trial by Bonomi and co-workers,[30] two dose levels of PAX were used – 135 mg/m^2 and 250 mg/m^2. Median survival was 9.5 and 10.1 months, respectively – a difference that was neither statistically nor clinically significant. The high-dose PAX group was given filgrastim, 5 µg/kg, for an unspecified number of days (making any costing problematic), to counteract bone marrow

toxicity leading to reduced numbers of white blood cells. There was no difference in the incidence of infection between high-dose and low-dose PAX. The low-dose group was not given filgrastim. At £73–117 per day, for perhaps 10–11 days, the routine use of filgrastim would add considerably to the cost of the regimen. However, because the higher dose of PAX conferred no significant advantage, the filgrastim cost is academic at present, until a trial of low-dose PAX with or without filgrastim is done.

Other drug-related costs

The SHPIC costing unit calculated intravenous administration at £24, reconstitution facilities (the costs of preparing cytotoxic drugs in a cytotoxic cabinet) at £5 and pharmacy staffing at £20. The minimal costs of diluent (normal saline taken with antiemetics) and disposables (needles, syringes, infusion bags) were ignored. It is assumed that staff costs for treatment delivery and monitoring are included in inpatient and outpatient costs. Total cost inflated to 1999/2000 prices is £52 per patient per administration.

SHPIC also reported that 1 hour of an NHS E-grade nurse's time would be required for prechemotherapy counselling. Taken as an average of the current pay scale and including national insurance and superannuation, this cost is calculated as £10.97.

Inpatient and outpatient costs

Few of the clinical trials or economic studies gave clear indications about requirements in terms of inpatient stay or outpatient visits for administration of chemotherapy (appendix 16), or admissions for chemotherapy-induced side-effects (appendix 17). In the latter case, studies either reported the percentages of patients admitted (without regard to duration), or gave admission to hospital a total cost (in the economic studies) without unit costs being available. The main cost of nausea and vomiting would arise if it were sufficiently severe to require that patients stayed in hospital overnight after chemotherapy. Jaakimainen and co-workers (1990)[15] reported hospitalisation rates, but based on low numbers of patients and on only two regimens (none of those considered here) and BSC. The lack of good published data on the requirements for admissions for administration of chemotherapy and for side-effects thereof creates uncertainty in attempting to model the relative cost-effectiveness of the proposed regimens. Nevertheless, estimates for chemotherapy administration were extracted from the data (supplemented by data from industry in one case).

Industry evidence suggested that the GEM+CDDP and VNB+CDDP regimens were administered on an outpatient basis in the trial by Comella and colleagues,[50] and expert opinion noted that PAX and DOC had also been administered on an outpatient basis.

One problem here is that, without better data, it is difficult to estimate total costs of chemotherapy, including not just the administration but the cost of side-effects. In the absence of available specific cost data on adverse events, we have applied a figure of £500, based on the estimate of one of our expert advisers, to each of the drug regimens to account for such things as admission for drug-induced neutropenia. Because of the lack of data, this figure has been applied irrespective of regimen, though it may vary among the four drugs, with the cost for the taxanes perhaps being higher.

The SHPIC costing unit has provided the costs of inpatient stay. These were calculated as the cost per day of stay in a respiratory medicine ward. (At Dundee Teaching Hospital Trust, patients with lung cancer were treated at King's Cross Hospital in the respiratory medicine ward, not in an oncology ward.) SHPIC calculated this inpatient cost at £132, incorporating overheads and direct costs (see appendix 18 for breakdown). The cost of an outpatient visit, assumed to be a medical consultant outpatient clinic, has been estimated from the Scottish Health Services Costs 'blue book' at £57. This value is the mean of the functional classification 01 and 02 hospitals (the large and general teaching hospitals in Scotland). Updating these hospital costs to 1999/2000 prices inflates an inpatient cost per day to £141 and an outpatient cost to £61 per visit. We were unable to find any published English reference costs. However, discussions with the Finance Department at Southampton General Hospital confirmed this figure to be similar to their costs of an outpatient visit.

'Best supportive care'

The term BSC is used to describe care that includes relief of symptoms by, for example, analgesics, but which does not attempt to prolong life or to remove (even if only temporarily) the cause of the symptoms. The term is useful to indicate the baseline option, but may vary in its inclusions. For example, radiotherapy may be part of palliative care, by providing temporary relief of metastatic symptoms.

Canadian BSC practice has been estimated by Evans[55,61,72] to cost Can$20,914 per patient (1993

prices). This estimate includes diagnostic and staging tests, hospitalisation for diagnosis and terminal care, palliative radiotherapy and clinic costs. As is apparent, Canadian BSC costs used for these calculations are significantly higher than Scottish BSC costs. Jaakimainen and co-workers[15] calculated BSC costs at Can$7236 (in 1984 prices), though this calculation was based on case notes for a small number of patients. An explanation for the significantly higher Canadian costs has been made elsewhere.[73]

The SHPIC costing unit calculated BSC costs in 1997/1998 based on data from the case notes of 36 patients with stage IV NSCLC at Dundee Teaching Hospital Trust. This was the only known UK calculation of BSC. None of the patients received chemotherapy. Cedian survival of these patients was 3.48 months. The costs incurred by these patients were averaged and found to be £4470 (detailed unit costs are given in appendix 19). It would have been desirable to calculate cost data from larger numbers of patients, but this was not possible within the constraints of a rapid review. Indeed, to highlight the problems of lack of data, similar estimates based on the experiences of small numbers of patients must have been made in the literature.[15]

The average cost of inpatient terminal care for these patients was calculated at £1341 per patient (appendix 19). The SHPIC costing unit assumed terminal care costs with and without prior chemotherapy as being equal, hence these were calculated and subtracted from BSC costs. The same process has been followed here. Therefore, the cost of BSC without terminal care was £3129 (assuming other direct costs are allocated proportionally by the cost of inpatient stay), which inflated to 1999/2000 prices is £3342.

We were able to check costs using data from the Southeast Scotland audit of lung cancer care, with much larger numbers, which should give more robust data on the cost of chemotherapy-related complications and the current cost of BSC. Raw data on resource use (e.g. admissions, outpatient attendances, primary care contacts, radiotherapy and blood transfusions) were obtained for 368 patients with stage III/IV NSCLC. (Unfortunately, no subgroup analysis was possible by stage.) Not surprisingly for a study originating in 1996–1997, only five patients received the newer regimens considered by this report (two received GEM, two GEM+CDDP and one VNB). Hence, we were unable to obtain any cost of side-effects complications for the new drugs. Excluding these and patients receiving other chemotherapy regimens left us with 316 patients in the BSC/palliative care group. (Radiotherapy was considered to be only palliative if the dose was 60 Gy or less.) On average, patients had 20.92 inpatient days, attended 5.07 outpatient visits, had 0.72 home-visit hours (43.2 minutes) and received 2.97 fractions of radiotherapy. The number of inpatient days from the SESLS data was similar to those in the SHPIC calculation (20.92 vs 19 days, respectively).

Using our previous unit costs, the new BSC calculations for inpatient stay (£141) and outpatient visits (£61) give a total of £3259. Home visits were conducted by a general practitioner (GP), district nurse or practice nurse. Half of the visits were by GPs and one-quarter each by practice and district nurses. Therefore, each patient, on average, received 21.6 minutes of GP visits, 10.8 minutes from a district nurse and 10.8 minutes from a practice nurse. The cost of a district nurse's time is £44 per hour, the cost of a practice nurse is £27 per hour, and the cost per home-visit minute of GPs is £3.06, according to Netten and Curtis.[70] The average total cost of all home visits per patient is therefore £79.

The cost of radiotherapy treatment, excluded from the above ward costs, was obtained from the National Schedule of Reference Costs. The mean average cost for less than 4 fractions of radiotherapy (with simulator) was £198. This mean cost is similar to the cost per attendance/fraction of £54 calculated by SHPIC (inflated to current prices). Using the Reference Costs figure, the average cost per patient would be £198. A further cost is for blood transfusions, which were necessary for some patients receiving palliative radiotherapy. On average, patients required 0.45 units of blood (most did not require any). The cost of the blood product (obtained from the National Blood Service) to hospitals is £78 per unit (£35 for 0.45 unit).

BSC cost per patient then totals £3572. This is also likely to be a slight underestimate of BSC cost because the collected SESLS data spanned only 6 months; 31 patients (10%) survived beyond this period (and would therefore have incurred further costs not monitored by the study). We have classified patients as having BSC based on initial intention to treat, and this excludes those receiving surgery, chemotherapy or radiotherapy as their primary treatment with the aim of cure or prolongation of life, but includes some who had palliative radiotherapy.

Full details of the SESLS costs will be published elsewhere, and the data given here are provisional. The advantage of including them is that the SHPIC costs may reflect factors such as a historically high provision on beds and a length of stay that is affected by local factors such as a large rural catchment and urban deprivation. The SESLS costs are probably more similar to those in most parts of England and Wales. However, the difference is small.

Cycles and doses

Data on median cycles, doses (in mg/m^2) and number of patients per study arm were extracted from each Phase III trial and certain robust Phase II trials. All these data were collated by drug regimen and are displayed in *Table 9*.[14,21–24,26–32,34,35,40,41,44,45,47,48,50,74,75]

In Models 2 and 3, we have had to aggregate doses and median cycles, unlike the practice in Model 1, which uses actual practice in the trials for dose and cycles. Here the aggregated published data provide the backbone for Models 2 and 3. Baseline, best and worst estimates (if different) have been drawn from *Table 9*. When a range of findings is given, baseline has followed the majority, in terms of median cycles and dose, or the study with the largest sample size. These data are tabulated in *Table 10*.

One problem is the definition of the number of administrations per cycle for the VNB regimens. To illustrate, in the case of GEM this is fairly straightforward. GEM is commonly given on days 1, 8 and 15 of each 28-day cycle. Therefore, each cycle consists of three administrations of GEM. With VNB, however, the situation appears more complex. Both the literature and recommendations state VNB is given weekly. The question is then, how long is a cycle? Clearly, if a cycle were 28 days, there would be four administrations of VNB per cycle; if a cycle were 21 days, there would be three administrations per cycle. Published data revealed in the trials were unclear here. However, Pierre Fabre Ltd provided supplementary evidence. The two largest VNB trials (Le Chevalier and co-workers[74] and Wozniak and colleagues[47]) gave the median number of cycles as 3 (*Table 9*). The Pierre Fabre Ltd data showed that the median number of **administrations** in each of these trials was 11 and 12, respectively, which suggests approximately four administrations per cycle of VNB. However, another company asserted that the VNB+CDDP arm in the (smaller VNB arm) trial by Comella and co-workers[50] used VNB weekly for 10 weeks, defined as 2 cycles, implying five administrations per cycle.

The Medical Research Council have reportedly adopted VNB regimens of 30 mg/m^2 on days 1 and 8 of each 3-week cycle (i.e. two administrations per cycle). We have followed the practice of the largest (Le Chevalier[74] and Wozniak[47]) trials in our baseline modelling, except for the pairwise comparisons in Model 1, for which we have followed the actual number of administrations stated (in the case of the trials by Comella and co-workers[50] and Le Chevalier and colleagues[74]). Cycles of PAX are based on a 3-week or 21-day cycle, with the average number of cycles most likely being 4 or 5.

However, the question remains whether one cycle of one regimen is equivalent to one cycle of another. Is it valid to compare regimens without comparing them with equal numbers of cycles? The answer to this is unclear, but an attempt has been made to address it in the sensitivity analysis.

Efficacy

Efficacy was analysed in terms of median survival rather than response, which some studies have used, because response is not necessarily indicative of increased length of life. It was not possible to present results by disease stage, given the lack of subgroup analysis in the reporting of survival data, though the majority of patients were at stage IV. In the case of median survival, a meta-analysis was not possible, but in Model 3 an attempt was made to calculate an overall effect for the economic analysis by combining the data (rather than anchor the model to one particular study's data).

In Models 1 and 3, we were forced to use median survival because that is what is reported in the literature. Only Berthelot and co-workers[54] recalculated mean survival from the raw data reported in trials, which is obviously beyond the scope of this review. In all cases, mean LYS was higher than medians, though the least difference was for BSC. Mean data from Berthelot and co-workers[54] have been examined (when applicable) in our sensitivity analysis, but we have retained the use of medians in our baseline models for a number of reasons: the inability to confirm the authors' data; mean survival is calculated for only some of the regimens we need to consider; their mean calculations comprise data from only one trial (a Phase II trial in the case of GEM); and in their recalculations of the raw data, they have considered only stage IV patients, an approach inconsistent with the rest of our model. Given the mean calculations all being greater than the medians, the effect of introducing mean data from Berthelot and co-workers[54] would be

TABLE 9 *Data extracted from trials and used in the models*

Drug regimen	Median no. of cycles	Median survival (range)	Dose (mg/m^2)	Study	*n* (study arm)	Jadad score
GEM	2	6.6 (4.9–7.3) months	1000	Bokkel-Huinink *et al.*, 1999[24]	72	2
GEM	2	6.6 (4.9–7.1) months	1000	Manegold *et al.*, 1997[75]	71	2
GEM	5	37 weeks or 8.5 months	1250	Perng *et al.*, 1997[28]	27	3
GEM+BSC	3	5.7 (4.6–7.6) months	1000	Anderson *et al.*, 2000[14]	150	3
GEM+CDDP	5	8.7 (7.7–10.2) months	GEM, 1250 CDDP, 100	Cardenal *et al.*, 1999[26]	69	2
GEM+CDDP	2 (70%), 4 (30%)*	42 weeks or 8.1 months	GEM, 1000 CDDP, 100	Comella *et al.*, 2000[50]	60	3
GEM+CDDP		8.6 months	GEM, 1000 CDDP, 100	Crino *et al.*, 1999[27]	155	2
GEM+CDDP	4	9.1 (8.3–10.6) months	GEM, 1000 CDDP, 100	Sandler *et al.*, 2000[29]	260	1
VNB		30 weeks or 6.9 months	25	Lorusso *et al.*, 1995[44]	35	2
VNB	9 administrations	30 weeks or 6.9 months	30	Crawford *et al.*, 1996[40]	143	3
VNB		32 weeks or 7.4 months	30	Depierre *et al.*, 1994[41]	119	3
VNB	3 (11 administrations)*	31 weeks or 7.2 months	30	Le Chevalier *et al.*, 1994[74]	206	3
VNB	4	28 weeks or 6.5 months	30	ELVIS Group, 1999[48]	76	3
VNB+CDDP		38 weeks or 8.8 months	VNB, 25 CDDP, 80	Lorusso *et al.*, 1995[44]	34	2
VNB+CDDP	2 (75%), 4 (25%) administrations/ cycle*	35 weeks or 6.7 months	VNB, 30 CDDP, 120	Comella *et al.*, 2000[50]	60	3
VNB+CDDP		33 weeks or 7.6 months	VNB, 30 CDDP, 80	Depierre *et al.*, 1994[41]	121	3
VNB+CDDP	3 (11 administrations)*	40 weeks or 9.2 months	VNB, 30 CDDP, 120	Le Chevalier *et al.*, 1994[74]	206	3
VNB+CDDP	4	9.6 (8.4–10.8) months	VNB, 25 CDDP, 60	Martoni *et al.*, 1998[45]	116	2
VNB+CDDP	3 (12 administrations)*	8 months	VNB, 25 CDDP, 100	Wozniak *et al.*, 1998[47]	206	2
PAX		24.1 weeks or 5.6 months	250	Chang *et al.*, 1993[31]	25	2
PAX+BSC	5	6.8 (5.7–10.2) months	200	Ranson *et al.*, 2000[32]	79	3
PAX+CDDP	6	9.7 months	PAX, 175 CDDP, 80	Giaccone *et al.*, 1998[35]	166	3
PAX(250)+CDDP	4	10 months	PAX, 250 CDDP, 75	Bonomi *et al.*, 2000[30]	191	2
PAX(135)+CDDP	5	9.4 months	PAX, 135 CDDP, 75	Bonomi *et al.*, 2000[30]	190	2
PAX+CDDP	5	8.1 months	PAX, 175 CDDP, 80	Gatzemeier *et al.*, 2000[34]	207	

continued

TABLE 9 contd *Data extracted from trials and used in the models*

Drug regimen	Median no. of cycles	Median survival (range)	Dose (mg/m^2)	Study	*n* (study arm)	Jadad score
DOC	3	6.0 (5.0–8.0) months	100	Roszkowski *et al.*, 2000[22]	137	2
DOC (second-line)		7 months	100 and 75	Shepherd *et al.*, 2000[21]	104	2
DOC (second-line)		5.9 months	100	Shepherd *et al.*, 2000[21]	49	2
DOC (second-line)		7.5 months	75	Shepherd *et al.*, 2000[21]	55	2
DOC (second-line)	3	5.5 months	100	Fossella *et al.*, 2000[23]	125	2
DOC (second-line)	3	5.7 months	75	Fossella *et al.*, 2000[23]	125	2
DOC+CDDP		None				
BSC		4.6 (3.7–6.0) months		Shepherd *et al.*, 2000[21]	100	2
BSC		5.7 (4.4–6.8) months		Roszkowski *et al.*, 2000[22]	70	2
BSC		5.9 (5.0–7.9) months		Anderson *et al.*, 2000[14]	150	3
BSC		4.8 (3.7–6.8) months		Ranson *et al.*, 2000[32]	78	3
BSC		21 weeks or 4.8 months		ELVIS Group, 1999[48]	78	3

* *Data provided by manufacturer (Pierre Fabre Ltd)*

to **increase** the cost-effectiveness of the regimens in the modelling.

Model 3 pooled median survival results from trials using the data in *Table 9*. Results by regimen were aggregated by patient numbers, with larger trials thus carrying more weight. Best and worst estimates, if different, were then defined by the upper and lower bounds of individual trial data. These data are collated in *Table 10*. Numbers of patients were used to weight the data rather than Jadad scores, given the lack of variation of these scores and the problems related to blinding in such circumstances detailed above. In the case of PAX, when doses (and thus drug cost) varied markedly between studies, this approach was not taken. Instead, several strategies were examined using the different PAX doses.

It is recognised this method of pooling the data consists of indirect comparisons between trials and is thus open to confounding. There may be differences in patient populations among trials. The control or comparator groups vary markedly between trials, thus we chose median survival rather than incremental survival. However, the mixture of different patient types may strengthen the conclusions and generalisability of the model. It is accepted that this method is not the ideal way of directly comparing regimens, but it does make the most of the data available, illustrates a range of possible cost-effectiveness estimates across a range of assumptions and can be interpreted with the aid of sensitivity analyses. (The underlying model will be made available with the version of this report on the NCCHTA website [http://www.ncchta.org] in order to allow others to do any other sensitivity analyses they may wish.)

Five studies reported median survival for BSC (*Table 9*). The average of these (also weighted by numbers of patients) was the figure used in the baseline analysis. However, the BSC regimens described in the studies permitted palliative treatment to varying degrees. For example, while palliative radiotherapy was always permitted, palliative chemotherapy was sometimes permitted. Commonly, decisions on treatment were left to individual physicians. The descriptions are presented in appendix 20. The effect of these variations is likely to affect quality of life and survival. A decision was made to combine the BSC estimates because BSC in practice will vary by consultant and hospital practice, and is independent of chemotherapy regimens used in other arms of a study; doing so should add to the generalisability of the model.

Survival data were also reported in terms of 1- and 2-year survival data in certain trials. The new drugs have given better survival than the BSC survival rates of 14–22% at 1 year and 0–7% at 2 years. Anderson and co-workers[14] reported GEM to give a 1-year survival of 25% and 2-year survival of 6%. One-year survival with GEM+CDDP has ranged from 32% to 39%, and with VNB from 25% to 32%. VNB+CDDP has shown 1-year survival of

TABLE 10 *Drug regimen values used in construction of the models*

		Estimate			Source
		Best	**Baseline**	**Worst**	
Drug dose (mg/m^2) in Models 2 and 3					
GEM			1000	1250	Derived from *Table 9*
CDDP+GEM			100		
VNB		25	30		
CDDP (with VNB)		60	120		
PAX			200	250	
PAX(135)+CDDP			135		
PAX(175)+CDDP			175		
PAX(250)+CDDP			250		
CDDP+PAX			75	80	
DOC			100		
DOC (second-line)			75	100	
Median number of drug cycles in Models 2 and 3					
GEM		2	3	5	Derived from *Table 9*
GEM+CDDP			4	5	
VNB			3	4	
VNB+CDDP			3	4	
PAX			5		
PAX(135)+CDDP			5		
PAX(175)+CDDP			5	6	
PAX(250)+CDDP			4		
DOC			3		
DOC (second-line)			3		
IP days or OP visits per cycle					
GEM	IP		0		Estimated (see appendices 15 and 17)
	OP		3		
GEM+CDDP	IP	0*	1		
	OP	3*	2		
VNB	IP		0		
	OP		4		
VNB+CDDP	IP	0*	1		
	OP	4*	3		
PAX	IP	0	1		
	OP	1	0		Expert reviewers†
PAX (any dose)+CDDP	IP		1		
	OP		0		
DOC	IP	0	1		
	OP	1	0		Expert reviewers†
DOC (second-line)	IP	0	1		
	OP	1	0		

continued

TABLE 10 contd *Drug regimen values used in construction of the models*

	Estimate			Source
	Best	**Baseline**	**Worst**	
Efficacy: median survival (months) in Model 3 only				
GEM	8.5	6.9	5.7	Derived from weighted average of data from *Table 9* (by numbers of patients)
GEM+CDDP	9.1	8.8	8.1	
VNB	7.4	7.1	6.5	
VNB+CDDP	9.6	8.4	6.7	
PAX	6.8	6.5	5.6	
PAX(135)+CDDP		9.4		
PAX(175)+CDDP	9.7	8.8	8.1	
PAX(250)+CDDP		10.0		
DOC	8.0	6.0	5.0	
DOC (second-line)	7.5	5.9	5.5	
BSC	4.6	5.2	5.9	

IP, inpatient; OP, outpatient

* *Additional data provided by manufacturer from the trial by Comella* et al., *2000*[50]

† *OP administration for PAX/DOC regimens was suggested by two expert reviewers*

36–39% and 2-year survival of 8–12%. Chang and colleagues[31] reported a 1-year survival of 42% with PAX, while 1- and 2-year survival rates of 37–43% and 19%, respectively, have been reported with PAX+CDDP. Roszkowski and co-workers[22] reported a 1-year survival of 25% and a 2-year survival of 12% with first-line DOC.

One study[49] used a GEM/VNB combination together with CDDP. However, the comparator was an unusual mixture (CDDP+EPI+VDS+LON), and the numbers of patients were small. For these reasons, and this being the only paper to date to consider this regimen, it has not been included in the model. Nor was a study comparing MITO+VNB+CDDP with MITO+VDS+CDDP[66] included in the analysis.

Quality of life

It is important to take account of QoL because we are dealing with small gains in quantity of life. The review of the clinical lung cancer studies shows a number of different QoL measures being used, if one is used at all. A number of studies simply offer a detailed assessment of adverse experiences or cancer-related symptom relief. Others may offer a monitoring of 'QoL' through performance status and symptom monitoring. Those papers that do use a QoL tool use the EORTC QLQ-C30, LC13 or LCSS questionnaires. One study of GEM monitored health-related QoL (HRQoL)[29] with the FACT-L questionnaire.

It would seem that there is as yet little consensus for obtaining QoL data from lung cancer trials, making it difficult to derive cost per QALY. Other options in a cost-effectiveness analysis are either cost per tumour response or cost per life-year gained (LYG). A cost per tumour response does not give as good an estimate of the cost-effectiveness of a drug as the cost per LYG. However, there are problems in simply doing a cost per LYG ICER, because the quality of an individual's life in his/her last remaining weeks may be of great importance.

The economic studies identified do not shed much more light on this problem. These papers either use cost-effectiveness ratios based on tumour response rates or on survival rates, or simply just assume the effectiveness to be equal between both treatment arms.

Two published studies have attached utility values to the chemotherapy regimens, but unfortunately neither of these has incorporated patient-based estimates. Smith and colleagues[62] used utility values of 0.7 for VNB and 0.6 for CDDP-containing

regimens in their QoL adjustments, which were based on estimates by 14 experienced oncology physicians and nurses. Similarly, a newly published study[54] based its utility weights on the views of 24 oncologists. These utility estimates were 0.60 for VNB and VNB+CDDP, 0.65 for GEM and 0.63 for PAX+CDDP, all higher than the 0.53 attributed to BSC. The estimates, while not adopted for our baseline model, have been considered in the sensitivity analysis.

Instead, we have considered qualitative information on the relative QoL impact of the regimens considered against BSC or other comparators from the trials.

Gemcitabine

One RCT reported QoL for GEM versus BSC.[14] This RCT used the EORTC QLQ-C30 and LC13 scales, and also the SS14 symptom scale. The EORTC scale was used at 2 and 4 months to test 25 variables. At 2 months, of the 25 variables tested, six showed an improvement greater than or equal to 10%; five variables improved for the GEM arm, with only one (dyspnoea) improving for the BSC arm. Five of the 25 variables showed a 10% or greater decrease in QoL, two in the GEM arm and three in the BSC arm. At 4 months, six of 25 variables showed a 10% or greater increase in QoL, all six being from the GEM arm. Of four variables that showed a deterioration of 10% or more, three were in the BSC arm, with only hair loss showing a deterioration in the GEM arm. Mean SS14 score from baseline to 2 months demonstrated a 10% improvement in symptoms with GEM and a 1% deterioration for BSC. A sustained improvement was achieved by 38% of GEM patients and 24% of BSC patients at 2 months, followed by 44% in the GEM arm and 25% in the BSC arm at 4 months, and 31% in the GEM arm and 22% in the BSC arm at 6 months.

One RCT comparing GEM versus CDDP+VP-16 reported QoL.[24] Using the EORTC QLQ-C30 and LC13 scales, the study found no statistically significant differences in change from mean baseline to after 6 cycles of treatment, for global QoL or functional scales. The only statistically significant difference seen in change from mean baseline scores was for alopecia, which worsened for CDDP+VP-16 after all cycles.

An RCT comparing GEM+CDDP and CDDP+VP-16 also used the EORTC-QLQC30 and LC13 scales.[26] It found no clinically significant differences in change from baseline within or between treatment arms in functional domains or global QoL. There was a significant improvement for both groups in pain, insomnia, cough, haemoptysis, chest pain and shoulder pain; however, there was no improvement in dyspnoea or fatigue for either arm. Peripheral neuropathy did not worsen in either group, but both had a significant worsening in nausea and alopecia. The only significant difference was in alopecia, which was worse for the CDDP+VP-16 arm.

Crino and co-workers (1999)[27] compared GEM+CDDP with MITO+IFOS+CDDP and reported QoL using the EORTC QLQ-C30 and LC13 scales. They found no overall difference in the change in QoL between the arms; the only differences were a worsening of alopecia from baseline in the MITO+IFOS+CDDP arm and a greater improvement in chest pain in the GEM+CDDP arm. Global QoL did not change significantly in either arm. There was a moderate decrease of physical functioning and a worsening of fatigue and nausea/vomiting. However, there was an improvement in pain, insomnia and cough.

Sandler and colleagues (2000)[29] compared GEM+CDDP with CDDP using the FACT-L HRQoL scale. There was no significant difference between the baseline and median change, at the last observation, between the two arms.

Conclusion

It appears that GEM improves the QoL compared with BSC. When compared with the other drugs used, GEM appears to have similar side-effects, except that alopecia may be less of a problem.

Vinorelbine

The Elderly Lung Cancer Vinorelbine Italian Study (ELVIS) Group (1999)[48] used the EORTC QLQ-C30 and LC13 questionnaires to ascertain the difference in QoL for VNB and BSC patients. At baseline, both groups of patients had the same QoL score, but the EORTC functional scales were consistently better, at each treatment cycle, for the VNB patients than the BSC patients, although the difference was statistically significant only for cognitive function and global health status. Among the symptom scales and items, VNB patients did better than the BSC patients for some lung-cancer-specific items (such as pain and dyspnoea), but worse for some treatment-related items (such as constipation, nausea/vomiting, hair loss and peripheral neuropathy).

Only two other RCTs report QoL results. One RCT[45] comparing VNB+CDDP with EPI+CDDP

reported QoL by monitoring performance status and symptoms, while the other[40] comparing VNB alone with 5FU+LV reported QoL using physician assessment (performance status and weekly symptom assessment) and patient assessment (SWOG questionnaire).

Martoni and co-workers (1998)[45] reported a greater than 10% improvement from baseline in performance status for both arms, although the difference was not significant. Crawford and colleagues (1996)[40] reported a non-significant improvement in symptom distress scores for VNB patients. There was an improvement in physical functioning for 5FU+LV patients. Some improvement in mean global QoL was reported for both groups. There was no difference in relief from cancer symptoms between treatment groups.

Conclusion
The one study directly comparing VNB with BSC shows some advantage for VNB in terms of its effect on QoL. The best we can do based on the other studies is note that VNB appears to improve symptoms compared with untreated baseline states.

Paclitaxel
Of the five RCTs reporting PAX compared with other treatments, one RCT[32] reported PAX versus BSC. This RCT used the RSCL as the QoL tool. Of the five RSCL subscores, the PAX arm showed a significant improvement in functional activity (impacts on functional activity were –0.094 for BSC and –0.091 with PAX, an absolute improvement in QoL of 0.003). All other subscores showed no significant difference.

Only two other RCTs report QoL. Bonomi and co-workers (2000)[30] compared PAX(two doses)+ CDDP versus CDDP+VP-16 and used the FACT-L scale to report QoL results. Giaccone and colleagues (1998)[35] compared PAX+CDDP versus CDDP+VM-26 and used the EORTC QLQ-C30-LC13 scale.

Bonomi and co-workers[30] showed that a higher percentage of patients receiving PAX+CDDP had an increased QoL compared with patients on CDDP+VP-16, but the difference was not significant because patients not completing the 6-week FACT-L assessment were excluded (19.7% vs 11.3%, respectively, an absolute difference of 8.3%). However, if these excluded patients were included, then there was a significant (8.8% absolute difference) improvement in QoL for patients receiving PAX+CDDP.

Giaccone and colleagues[35] reported a significant improvement for the PAX+CDDP arm in most functioning scales at the 6-week assessment. On the symptom scale, there was a significant decrease in fatigue and appetite loss at the 6-week assessment for the PAX+CDDP arm (absolute difference not stated in the paper). However, the benefit was no longer significant at the 12-week assessment. There was an increase in peripheral neuropathy for the PAX+CDDP arm at the 12-week assessment.

Conclusion
The one trial that compared PAX with BSC shows that, at worst, PAX causes no reduction in QoL, and it appears to show marginal benefit in one of the five subscores.

Docetaxel
There are two papers[21,22] looking at the effect of DOC versus BSC in NSCLC. Both these studies measured QoL. One trial[21] used both the LCSS, for North American patients, and the EORTC scale, for European patients, to assess QoL. The other used only the EORTC scale.[22] The QoL results favour the patients receiving DOC in both studies. There was a significant improvement in pain and fatigue symptoms for patients on DOC.[21] Emotional functioning was significantly improved for patients on DOC,[22] as were nausea/vomiting, pain and dyspnoea. The only dimension not to produce a favourable result for the DOC arm was diarrhoea.[22] Roszkowski and co-workers (2000)[22] referred to the use of DOC as first-line chemotherapy, while Shepherd and colleagues (2000)[21] considered its use only as second-line therapy, so the type of BSC provided in these two studies may not be the same.

Conclusion
DOC appears to improve QoL compared with BSC in both second-line and first-line therapy.

Direct comparisons of the four regimens
Unfortunately, no RCTs directly compared two or more of the regimens in terms of QoL, thus side-by-side comparisons cannot be made.

Comella and co-workers (1999)[49] reported the QoL of patients treated using a combination of GEM and VNB. A 10-item LCSS-derived questionnaire was used; however, this is of little use in comparisons because GEM and VNB were incorporated in the same regimen. They compare CDDP+GEM+VNB with CDDP+EPI+VDS+LON. This RCT reported an improvement over baseline QoL for 59% of CDDP+GEM+VNB patients, while reporting an improvement in QoL for 39% of

CDDP+EPI+VDS+LON patients. No significance levels were reported.

Overall conclusion

It is worth highlighting the modest survival gains achieved by chemotherapy. There would be little benefit in a few extra months of life if they were made miserable by drug side-effects. However, the key finding from the studies that reported QoL is that, at worst, the drugs do not reduce overall QoL and that, in some instances, they improve it, when compared with BSC. Uncontrolled metastases also have 'side-effects'.

Adverse events

All the drugs used in lung cancer chemotherapy have side-effects. In the absence of head-to-head trials, any comparison of side-effect frequency or severity has to be indirect. Appendix 21 summarises the frequency of side-effects, as reported in the studies. The CDDP-containing regimens appear to have higher levels of severe adverse effects, while GEM and VNB arguably have the least. Two studies found very different rates (4% and 53%) of neutropenia with VNB.[43,48] The low estimate in one study was unusually low and may relate to the completeness of safety data in this study, which concentrated on QoL measurement (see appendix 21).

Sensitivity analysis

One-way sensitivity analysis was carried out across a range of variables in each of the models. The first sequence concerns the variability of regimen cycles: assuming 3 cycles for each regimen; assuming 60% of patients had only 1.5 cycles, while 40% went on to complete 3 cycles (suggested by experts and arguably a more realistic scenario because not all patients will want chemotherapy or be suitable for a full course); the number of administrations per cycle of VNB, as discussed above (page 41); the best and worst cycles indicated in *Table 10*. An industry submission (Bristol-Myers Squibb) gave the mean number of cycles for the PAX-containing regimens as 4, and this number of cycles has also been evaluated here. Next we consider the effect of drug costs that are lower than published in the BNF, a scenario that may arise with the availability of generic products or when drug companies reduce the drug cost to increase the cost-effectiveness of their product. Two rates of discount were considered: 25% and 50%. (The CDDP components were not discounted.) An estimate of £150 to account for the cost of the newer antiemetics regimens was also included. The mean survival estimates calculated by Berthelot and co-workers[54] were used, if applicable, in Models 1 and 3. Similarly, the Berthelot report's[54] utility estimates for QoL were examined. The outpatient administration applicability for certain regimens indi-cated in *Table 10* was used in each of the models. In Model 3, the best and worst survival estimates from *Table 10* were used. The SESLS BSC estimate was included, as was a lower BSC estimate of £2200 (approximately one-third less than the SHPIC figure). The reduced dose scenario described above in the *Drug costs* section has also been considered here (though affecting only the VNB regimens, as stated earlier).

Results of economic analysis

Model 1: direct pairwise comparisons of the newer regimens or with BSC

Eight trials provided enough data on pairwise comparisons of the newer regimens or with BSC to model their cost-effectiveness. The following model, shown in *Table 11*, uses in its calculations the doses, cycles and administration of the relevant individual trial from *Table 9*. Each of the newer regimens is compared with BSC, with only two trials directly comparing the newer drugs. Comella and co-workers[50] compared VNB+CDDP against GEM+CDDP, and Le Chevalier and colleagues[74] compared VNB with VNB+CDDP. Finally, two separate doses of second-line DOC (used after the failure of previous chemotherapy) were compared with BSC (not strictly a pairwise comparison) and each other twice. Further direct comparisons of the newer regimens are clearly needed. Cost-effectiveness is given as cost per LYS versus the respective comparator.

Results

The trial by Anderson and co-workers[14] showed GEM dominated by BSC across a range of assumptions because of the median survival in the BSC arm. However, concern has been raised over the large proportion of patients who received palliative radiotherapy in this trial (79% in BSC arm vs 49% in GEM arm). Interestingly, when the Berthelot report's[54] QoL utilities are used, there is a gain of 0.05 QALYs per patient despite the lower LYS, and the incremental cost per QALY with GEM is £16,388. The ELVIS Group[48] has shown VNB to dominate BSC in terms of lower cost and longer survival. This dominating effect or reasonable cost-effectiveness is maintained across the range of sensitivity assumptions, though this was a small trial with only 76 and 78 patients in the VNB and BSC arms, respectively. The PAX and DOC (first-line) regimens look expensive compared with BSC in all but a

TABLE 11 *Model 1: direct pairwise comparisons between newer regimens or between newer regimens and BSC*

	Anderson *et al.*, 2000[14]		ELVIS Group, 1999[48]		Comella *et al.*, 2000[50]				Ranson *et al.*, 2000[32]		Roszkowski *et al.*, 2000[22]		Le Chevalier *et al.*, 1994[74]		Shepherd *et al.*, 2000[21]			Fossella *et al.*, 2000[23]	
	BSC	GEM	BSC	VNB	VNB+CDDP		GEM+CDDP		BSC	PAX	BSC	DOC	VNB	VNB+ CDDP	BSC	DOC 2L (75 mg/m^2)	DOC 2L (100 mg/m^2)	DOC 2L (75 mg/m^2)	DOC 2L (100 mg/m^2)
Median no. of cycles		3		4	2	4	2	4		5		3	3	3		3	3	3	3
Median no. of administrations (e.g. GEM and VNB)		9		8	10*	20*	6	12		5		3	11*	11*		3	3	3	3
Median no. of CDDP administrations					2*	4*	2	4						3					
Proportion of patients*					75%*	25%*	70%*	30%*											
Drug cost (e.g. GEM and VNB) (£)		2,637		1,426	1,337	892	1,230	1,055		6,858		3,975	1,961	1,961		3,300	3,975	3,300	3,975
Drug cost (CDDP) (£)					109	73	85	73						219					
Administration costs (£)		471		419		785		544		262		157	576	733		157	157	157	157
Counselling costs (£)		11		11		11		11		11		11	11	11		11	11	11	11
Inpatient administration (£)		0		0		0		0		662		397	0	397		397	397	397	397
Outpatient administration (£)		513		456		713*		445*		0		0	627	627		0	0	0	0
Side-effects (estimated) (£)		500		500		500		500		500		500	500	500		500	500	500	500
Total cost per patient (£)	3,342	4,132	3,342	2,812		4,420		3,943	3,342	8,293	3,342	5,040	3,675	4,448	3,342	4,365	5,040	4,365	5,040
Incremental cost (vs comparator) (£)		789		–530				–477		4,951		1,698		773		1,023	675		675
Median survival (months)	5.9	5.7	4.8	6.5		6.7		8.1	4.8	6.8	5.7	6.0	7.2	9.2	4.6	7.5	5.9	5.7	5.5
LYS	0.49	0.48	0.40	0.54		0.56		0.68	0.40	0.57	0.48	0.50	0.60	0.77	0.38	0.63	0.49	0.48	0.46
Average cost per LYS (£)	6,798	8,698	8,355	5,192		7,916		5,841	8,355	14,634	7,036	10,081	6,125	5,802	8,719	6,985	10,252	9,190	10,997
Incremental median survival (months)		–0.2		1.7				1.4		2.0		0.3		2.0		2.9	–1.6		–0.2
Incremental LYS		–0.02		0.14				0.12		0.17		0.03		0.17		0.24	–0.13		–0.02
Incremental cost per LYS (vs comparator) (£)		D–		D+				D+		29,704		67,926		4,638		4,234	D–		D–

continued

TABLE 11 contd *Model 1: direct pairwise comparisons between newer regimens or between newer regimens and BSC*

	Anderson *et al.*, 2000[14]		ELVIS Group, 1999[48]		Comella *et al.*, 2000[50]		Ranson *et al.*, 2000[32]		Roszkowski *et al.*, 2000[22]		Le Chevalier *et al.*, 1994[74]		Shepherd *et al.*, 2000[21]			Fossella *et al.*, 2000[23]	
	BSC	GEM	BSC	VNB	VNB+CDDP	GEM+CDDP	BSC	PAX	BSC	DOC	VNB	VNB+ CDDP	BSC	DOC 2L (75 mg/m^2)	DOC 2L (100 mg/m^2)	DOC 2L (75 mg/m^2)	DOC 2L (100 mg/m^2)
One-way sensitivity analysis: incremental cost per LYS (vs comparator) (£)																	
Baseline	D–		D+		D+		29,704		67,926		4,638		4,234		D–		D–
VNB cycle – 21 days/ 3 administrations	NA		4,380		NA		NA		NA		NA		NA		NA		NA
VNB cycle – 21 days/ 2 administrations (days 1 and 8)*	NA		NA		NA		NA		NA		NA		NA		NA		NA
25% discount on BNF	D–		D+		D+		19,417		28,176		4,638		820		D–		D–
50% discount on BNF	CS–		D+		D+		9,130		D+		4,638		D+		D–		D–
Antiemetics (£150)	D–		D+		D+		30,604		73,926		4,638		4,854		D–		D–
Berthelot's mean survival	NA		NA		NA		NA		NA		4,832		NA		NA		NA
Berthelot's QoL utilities	16,388		D+		NA		NA		NA		7,731		NA		NA		NA
Best survival	3,643		NA		NA		9,140		5,661		NA		NA		NA		NA
Worst survival	D–		NA		D+ (stage IV)		D–		D–		NA		NA		NA		NA
Lower BSC estimate (£2,200)	D–		4,231		NA		36,557		113,614		NA		8,960		D–		D–
SESLS BSC cost (£3,572)	D–		D+		NA		28,325		58,734		NA		3,283		D–		D–
Best cycles	NA		NA		NA		NA		NA		NA		NA		NA		NA
Worst cycles	NA		NA		NA		NA		NA		NA		NA		NA		NA
Outpatient administration (see *Table 10*)	NA		NA		D+		27,440		58,871		2,254		NA		NA		NA
Reduced dose (if less than 20% of vial used)*	NA		D+		D+		NA		NA				NA		NA		NA

2L, second-line therapy; D–, dominated by comparator (higher cost and lower survival than comparator); D+, dominant strategy vs comparator (it is not appropriate to show a figure when a strategy is dominant); NA, not applicable; CS–, cost-saving

* *Data provided by manufacturer*

couple of scenarios in each of the Ranson[32] and Roszkowski[22] trials.

A substantial amount of data on the administration of the regimens of the Comella trial[50] was provided by industry. In this trial, GEM+CDDP dominated VNB+CDDP across the range of sensitivity assumptions. (The range of VNB administrations per cycle has not been examined here because the number of administrations was explicitly stated.) However, a large number of concerns have been raised about this paper. A fourth arm (PAX+CDDP+GEM) was reportedly introduced part-way through the trial, and this is not mentioned in the *Journal of Clinical Oncology* article,[50] nor is its impact on the randomisation procedure. In any case, this is one small trial with each arm consisting of only 60 patients.

VNB+CDDP provides reasonable incremental cost per LYS compared with VNB alone through the range of assumptions, according to Le Chevalier and colleagues.[74]

The 100-mg/m^2 dose of DOC is dominated by the 75-mg/m^2 dose, which itself offers a reasonable incremental cost-effectiveness versus BSC.

Whatever the methodological quality of these trials, they are only single trials with small numbers of patients. More direct comparisons are needed.

Model 2: cost-minimisation analysis

Given the lack of side-by-side comparisons between the newer regimens, it is arguably inappropriate to make any definitive conclusion about relative efficacy. One way of dealing with this is to assume equal survival and conduct a cost-minimisation analysis. This is the approach taken for this model (*Table 12*). All regimens have BSC as their comparator treatment because this remains standard treatment for the majority of patients in the UK. Results are presented in terms of average cost per patient and incremental cost versus BSC. Number of cycles, doses and inpatient/outpatient administration are as in the aggregation in *Table 10*.

Results

The single agents VNB and GEM have the least incremental cost relative to BSC. This is maintained across the majority of sensitivity analysis assumptions. However, these single agents tend to lead to lower incremental survival than the CDDP-containing regimens. Then VNB+CDDP and DOC follow, with the PAX(135 mg/m^2)+CDDP and GEM+CDDP regimens being of slightly higher cost. The most expensive regimens are single-agent PAX, PAX(175 mg/m^2)+CDDP and PAX(250 mg/m^2)+CDDP. DOC for second-line therapy looks reasonably cost-effective compared with BSC across the range of sensitivity assumptions.

Assuming equivalence of cycles between the GEM and VNB regimens, and that 3 cycles of each were prescribed, the GEM+CDDP and PAX(135 mg/m^2)+CDDP regimens become more cost-effective versus BSC, and GEM+CDDP approaches the cost of VNB+CDDP. However, it is unclear if this is valid for the PAX-containing regimens, and a more realistic scenario is probably 4 cycles. The scenario for which there is a 60% probability of patients receiving 1.5 cycles and 40% probability of getting 3 cycles further favours the VNB, GEM and DOC regimens in terms of cost per patient. However, it is less clear, as noted above, whether this is a realistic scenario for the PAX-containing regimens.

Model 3: cost-effectiveness analysis

Due to the number of comparators, no meta-analysis directly combining the drugs was possible, as detailed in the *Effectiveness* chapter. This model uses the trial data as best we can to give a broad picture of likely relative cost-effectiveness compared with BSC (*Table 13*). Number of cycles, doses, inpatient/outpatient administration and median survival follow the aggregation in *Table 10*. We have included an 8-administration VNB regimen because this was the regimen actually received in the ELVIS study.[48] Again, all regimens have BSC as their comparator treatment because this remains standard treatment for the majority of patients in the UK. Results are presented in terms of incremental cost per LYS versus BSC. Caution should be used in any inter-regimen comparisons because of the way the data were combined.

Results

The regimens with the least incremental cost-effectiveness over BSC under the baseline scenario are VNB, VNB+CDDP and GEM. These regimens retain their cost-effectiveness under a range of assumptions and may even be dominant under certain circumstances. However, for GEM, the incremental cost per LYS rises to around £20,000 under the worst assumptions of cycles and survival. While most scenarios favour the VNB regimens, the GEM and VNB regimens deliver similar levels of cost-effectiveness if the same number of cycles are applied and the length of a cycle is deemed equivalent at 28 days (or four VNB administrations). The lower-dose PAX+CDDP regimens also offer favourable cost-effectiveness with 3 cycles, but there is no published evidence whether this is an appropriate scenario for the PAX regimens. The favourability of

TABLE 12 *Model 2: cost-minimisation analysis (assuming equal efficacy)*

	BSC	GEM	GEM+CDDP	VNB	VNB+CDDP	PAX	PAX (135 mg/m^2) +CDDP	PAX (175 mg/m^2) +CDDP	PAX (250 mg/m^2) +CDDP	DOC	DOC 2L
Median no. of cycles		3	4	3	3	5	5	5	4	3	3
No. of administrations (e.g. GEM and VNB)		9	12	11	12	5	5	5	4	3	3
No. of CCDP administrations		0	4	0	3	0	5	5	4	0	0
Drug cost (e.g. GEM and VNB) (£)		2637	3516	1961	2140	6858	4364	5610	6483	3975	3300
Drug cost (CDDP) (£)			243		219		243	243	204		
Administration costs (£)		471	837	576	785	262	523	523	419	157	157
Counselling costs (£)		11	11	11	11	11	11	11	11	11	11
Inpatient administration (£)		0	530	0	397	662	662	662	530	397	397
Outpatient administration (£)		513	684	627	684	0	0	0	0	0	0
Side-effects (estimated) (£)		500	500	500	500	500	500	500	500	500	500
Average cost per patient (£)	3342	4132	6321	3675	4736	8293	6304	7550	8147	5040	4365
Incremental cost (vs BSC) (£)		789	2979	333	1394	4951	2962	4208	4804	1698	1023
One-way sensitivity analysis: incremental cost (vs BSC) (£)											
Baseline		789	2979	333	1394	4951	2962	4208	4804	1698	1023
All regimens – 3 cycles		789	1526	333	1394	1838	645	1392	2896	1698	1023
60% of patients get 1.5 cycles, 40% get 3 cycles		CS	735	CS	752	603	CS	382	1437	506	33
PAX regimen – mean of 4 cycles*		NA	NA	NA	NA	3394	1804	2800	4804	NA	NA
VNB single agent – 8 administrations total		NA	NA	CS	NA	NA	NA	NA	NA	NA	NA
VNB cycle – 21 days/3 administrations		NA	NA	CS	531	NA	NA	NA	NA	NA	NA
VNB cycle – 21 days/2 administrations (days 1 and 8)*		NA	NA	CS	CS	NA	NA	NA	NA	NA	NA
25% discount on BNF (excluding CDDP)		130	2100	CS	859	3236	1871	2805	3184	704	198
50% discount on BNF (excluding CDDP)		CS	1221	CS	324	1522	780	1403	1563	CS	CS
Antiemetics (£150)		939	3129	483	1544	5101	3112	4358	4954	1848	1173
Lower BSC estimate (£2200)		1932	4121	1475	2536	6093	4104	5350	5947	2840	2165
SESLS BSC cost (£3572)		560	2749	103	1164	4721	2732	3978	4575	1468	793
Best cycles		CS	NA	NA	NA	NA	NA	NA	NA	NA	NA
Worst cycles		3203	4431	620	2802	NA	NA	5616	NA	NA	NA
Outpatient administration (see *Table 10*)		NA	2449	NA	996	4573	NA	NA	NA	1472	NA
Reduced dose (if less than 20% of vial used)*		NA	NA	CS	1019	NA	NA	NA	NA	NA	NA

2L, second-line therapy; CS, cost-saving vs BSC

* *Data provided by manufacturer*

TABLE 13 *Model 3: cost-effectiveness analysis*

	BSC	GEM	GEM+CDDP	VNB	VNB+CDDP	PAX	PAX (135 mg/m^2) +CDDP	PAX (175 mg/m^2) +CDDP	PAX (250 mg/m^2) +CDDP	DOC	DOC 2L
Median no. of cycles		3	4	3	3	5	5	5	4	3	3
No. of administrations (e.g. GEM and VNB)		9	12	11	12	5	5	5	4	3	3
No. of CCDP administrations		0	4	0	3	0	5	5	4	0	0
Drug cost (e.g. GEM and VNB) (£)		2,637	3,516	1,961	2,140	6,858	4,364	5,610	6,483	3,975	3,300
Drug cost (CDDP) (£)			243		219		243	243	204		
Administration costs (£)		471	837	576	785	262	523	523	419	157	157
Counselling costs (£)		11	11	11	11	11	11	11	11	11	11
Inpatient administration (£)		0	530	0	397	662	662	662	530	397	397
Outpatient administration (£)		513	684	627	684	0	0	0	0	0	0
Side-effects (estimated) (£)		500	500	500	500	500	500	500	500	500	500
Average cost per patient (£)	3,342	4,132	6,321	3,675	4,736	8,293	6,304	7,550	8,147	5,040	4,365
Incremental cost (vs BSC) (£)		789	2,979	333	1,394	4,951	2,962	4,208	4,804	1,698	1,023
Median survival (months)	5.24	6.90	8.80	7.06	8.45	6.51	9.40	8.81	10.00	6.00	5.94
LYS	0.44	0.58	0.73	0.59	0.70	0.54	0.78	0.73	0.83	0.50	0.49
Average cost per LYS (£)	7,658	7,184	8,623	6,249	6,726	15,283	8,048	10,281	9,776	10,081	8,824
Incremental median survival (months)		1.66	3.56	1.82	3.21	1.27	4.16	3.58	4.76	0.76	0.70
Incremental LYS		0.14	0.30	0.15	0.27	0.11	0.35	0.30	0.40	0.06	0.06
Incremental cost per LYS (vs BSC) (£)		5,690	10,041	2,194	5,206	46,610	8,537	14,124	12,104	26,707	17,546
One-way sensitivity analysis: incremental cost per LYS (vs BSC) (£)											
Baseline		5,690	10,041	2,194	5,206	46,610	8,537	14,124	12,104	26,707	17,546
All regimens – 3 cycles		5,690	5,145	2,194	5,206	17,304	1,858	4,673	7,295	26,707	17,546
60% of patients get 1.5 cycles, 40% get 3 cycles		D+	2,478	D+	2,808	5,682	D+	1,283	3,620	7,952	569
PAX regimens – mean of 4 cycles*		NA	NA	NA	NA	31,957	5,198	9,398	12,104	NA	NA
VNB single agent – 8 administrations total		NA	NA	D+	NA	NA	NA	NA	NA	NA	NA
VNB cycle – 21 days/3 administrations		NA	NA	D+	1,982	NA	NA	NA	NA	NA	NA
VNB cycle – 21 days/2 administrations (days 1 and 8)*		NA	NA	D+	D+	NA	NA	NA	NA	NA	NA
25% discount on BNF		939	7,079	D+	3,208	30,469	5,392	9,416	8,021	11,078	3,398
50% discount on BNF		D+	4,116	D+	1,209	14,327	2,248	4,708	3,937	D+	D+
Antiemetics (£150)		6,771	10,547	3,183	5,767	48,023	8,970	14,627	12,482	29,066	20,119
Berthelot's mean survival		1,925	NA	1,189	3,167	NA	5,196	NA	8,429	NA	NA
Berthelot's QoL utilities		5,538	NA	2,738	7,290	NA	11,296	NA	16,358	NA	NA
Best survival		2,903	9,253	1,846	3,833	38,008	NA	11,314	NA	7,375	5,425
Worst survival		20,458	12,484	3,162	11,430	163,647	NA	17,636	NA	D–	46,680
Lower BSC estimate (£2,200)		13,923	13,892	9,725	9,473	57,364	11,830	17,957	14,982	44,670	37,134
SESLS BSC cost (£3,572)		4,034	9,267	679	4,348	44,447	7,875	13,352	11,525	23,093	13,605
Best cycles		D+	NA	NA	NA	NA	NA	NA	NA	NA	NA
Worst cycles		23,089	14,938	2,194	10,467	NA	NA	18,849	NA	NA	NA
Outpatient administration (see *Table 10*)		NA	8,255	NA	3,722	43,058	NA	NA	NA	23,146	13,664
Reduced dose (if less than 20% of vial used)*		NA	NA	D+	3,805	NA	NA	NA	NA	NA	NA

2L, second-line therapy; D–, dominated by BSC (higher cost and lower survival than BSC); D+, dominant strategy vs BSC (it is not appropriate to show a figure when a strategy is dominant)

* *Data provided by manufacturer in industry submission*

the PAX+CDDP regimens is attributed to those regimens tending to have the highest incremental survival over BSC. The unlicensed PAX and DOC single agents are the most expensive compared with BSC in the baseline case, though it is worth noting the first-line DOC regimen came out favourably in the cost-minimisation analysis.

We repeat that, because of the limitations of the data and hence use of this method of pooling them, we would caution against direct comparisons between regimens. We considered a marginal analysis inappropriate. However, the results show the reasonable cost-effectiveness of GEM, GEM+CDDP, VNB, VNB+CDDP and the lowest-dose PAX+CDDP regimens, compared with BSC under a range of scenarios and assumptions.

DOC as second-line therapy appears to be relatively expensive in the baseline scenario due to its small survival gain over BSC, but would be used for only small numbers of patients. It is most cost-effective when assuming larger survival gains or discounts on the BNF price, but performs worst when BSC is less expensive and survival more closely approximates BSC.

Discussion of economic analysis

The main conclusion of the economic analysis is that chemotherapy for NSCLC is cost-effective, taking into account both survival and QoL.

Each of the above three models has a number of limitations. The cost-effectiveness estimates in the pairwise comparisons are based on single studies. The cost-minimisation analysis is of limited use in policy decision-making, especially when it is unlikely that the regimens have equal efficacy in practice. The cost-effectiveness analysis had to use an unorthodox method for pooling individual trial data. Despite this, the pairwise comparisons and the cost-effectiveness analysis do illustrate the reasonable levels of cost-effectiveness relative to BSC across a range of assumptions for the GEM, VNB and PAX+CDDP regimens. The ICERs shown in Model 3 for these regimens are favourable when compared with other healthcare interventions recently funded.

An alternative approach to pooling the efficacy data, through a more formal statistical method using standard errors for the difference in medians along the lines of Gore and co-workers (1992),[76] was suggested by one of our reviewers. However, this approach has not been possible within the resource constraints of this review.

Patient choice

Not all patients would wish to undergo therapy other than palliative treatment, if given opportunities for informed choice. For example, the survey of patients by Silvestri and colleagues (1998)[77] reported that the patients (n = 81) would not want chemotherapy unless median survival improved by 4.5 months for mild toxicity and 9.0 months for severe toxicity. However, one of our expert reviewers reported a lack of understanding by patients of the effects and side-effects of chemotherapy – a general belief of the patients that the side-effects of such treatments outweigh any benefits.

Raby and co-workers (1995)[78] conducted a postal survey in 1993 of Canadian respiratory physicians, thoracic surgeons, radiation oncologists and medical oncologists. For stage IIIB NSCLC, 17% recommended no treatment, 65% recommended radiotherapy alone, 2% recommended chemotherapy alone, and 16% recommended chemotherapy with radiotherapy. For stage IV NSCLC, 80% recommended no treatment, and 20% recommended chemotherapy. However, this survey was conducted before the new regimens considered in this report generally came onto the scene, and their lower toxicity compared with some older regimens used in 1993 will probably have altered clinical opinion.

Each of the CDDP-containing regimens gives increased survival, which must be traded off against higher costs and toxicity.

Cost-effectiveness

The cost-effectiveness data published by previous economic studies are of limited use in a UK NHS setting. Perhaps the Canadian Population Health Model (POHEM)[54] could be adapted, but this would require clarification of some items in the model by its designers, and replacement with UK data that are at present not readily available. The Canadian authors do agree with our conclusions that the regimens appear to be effective and relatively cost-effective.

The cost per LYG is used as the outcome because of the lack of good data on the relative effects on QoL of the different drugs. We have therefore not produced a cost per QALY, but given that chemotherapy seems to at least maintain and probably modestly improve QoL, adjusting for QoL would arguably make little difference and might improve the ICER. The illustration in the sensitivity analyses, using the Berthelot report's[54] utility estimates, supports this argument.

Other QoL issues

QoL will not remain constant over the survival period, and patients are likely to pass through a number of distinguishable health states. In SCLC, Rosenthal and co-workers (1992),[79] utilising quality-adjusted time without symptoms or toxicity (Q-TWiST), assigned utility values to different stages of chemotherapy: 1.0 for time without symptoms or toxicity (TWiST), 0.75 for toxicity, 0.5 for symptoms and 0.25 for relapse. However, these utility values were estimated and not patient based. It is questionable, for example, whether the TWiST period qualifies for a utility value of 1.0. The most valuable period of life for chemotherapy patients is that after the cessation of chemotherapy admission before relapse. This is when QoL will be highest. However, given the significant proportion of time spent receiving chemotherapy (4 or 5 months), the QoL of this period is also of concern. Rosenthal and co-workers[79] estimated that the QoL under chemotherapy administration (0.75) is higher than that for untreated symptoms (0.5). In their sensitivity analysis, they found that "the most influential utility coefficient was that for toxicity". This reflects the considerable time spent receiving treatment as a proportion of overall survival in these patients.

These estimates need to be supported by patient-based data observed at the various stages (though this may itself be biased in favour of the healthiest patients responding).

Are all QALYs equal?

The duration of life needs to be considered. The usual reason for expressing benefits in terms of cost per QALY is that it provides a common currency, to allow policy-makers to compare the relative benefits of different ways of allocating healthcare resources. However, this assumes that all QALYs are equal – that a 6-month QALY gain is always worth the same, whether it is 6 months added to 6 months, as is roughly the case with chemotherapy for NSCLC, or 6 months added to 5 or 10 years, as might be the case with some uses of statins.

It has been argued that the concept of diminishing marginal utility should be taken into account. If so, the 6-month gain that is a large part of what life expectancy would be left should be valued more highly than if the 6 months is a small part of a much longer period (i.e. that patients value a short extension to a short expectancy more than a short extension to a longer expectancy). How much more is for debate. Waugh and Scott (1998)[80] tried to provoke discussion on this through the correspondence column of the *BMJ*, and suggested an inverse weighting with doubling of life-year gain if total life expectancy was under 1 year. The main response supported the use of a diminishing marginal utility weighting.[81]

Antiemetics

As noted above (page 38), the cost of antiemetics was excluded from the baseline model due to their negligible cost. However, if these older drugs used in the trials were to be replaced by newer, more costly regimens including, for example, granisetron (£36 per 3 mg), costs would rise, but as shown in *Table 13,* the difference would be insufficient to affect the cost-effectiveness of the less expensive regimens.

Future developments with these drugs

The drug cost of PAX may well fall in the near future, with the removal of the patent and availability of generic versions. The US drug company Ivax plans to release a generic version of the Bristol-Myers Squibb drug by the end of October 2001 (*Wall Street Journal Europe*, 2000 Sep 8–9; *Financial Times*, 2000 Oct 16). While not achieving cost-saving status, substantial discounts are likely to lead to reasonable cost-effectiveness ratios compared with BSC, GEM and VNB+CDDP.

Oral versions of drugs would reduce administration costs. An oral formulation of VNB is currently being assessed by the Medicines Control Agency (MCA) and has already been approved in some countries. This oral version will reduce the costs of administration (no infusion will be needed, thereby reducing the workload on nursing and pharmacy) and the number of visits needed, though specialist monitoring of response and side-effects will still be needed. The net effect will be to increase the cost-effectiveness of VNB and reduce inconvenience to patients.

Administration of paclitaxel on an outpatient basis

PAX has been given successfully on an outpatient basis in Canada (Coyle D, The Ottawa Hospital, Ottawa, Canada: personal communication, 2001). However, substituting outpatient visits for inpatient stay would have little impact on the cost per LYS for the single agent (*Table 12*), other than freeing up hospital beds for alternative uses.

Carboplatin rather than cisplatin

It is necessary for patients receiving CDDP to have an intravenous infusion of fluid for renal protection. This is less of a problem with CBDCA, which could be given on an outpatient basis.

Admission for side-effects

Owing to the lack of data on admissions for chemotherapy-induced side-effects, an estimate for this has been entered into the model. Were significant additional stays included, these would compromise the cost-effectiveness of the regimens. The same amount has been entered for all the new drugs, but there is probably variation among them. In the absence of good data, an estimate had to be used and was based on expert advice. This estimate was somewhat higher than one provided in a confidential industry submission. However, it may be noted that, despite an increase in the use of chemotherapy, the average number of days spent in hospital between diagnosis and death has been falling. While this decrease would be compatible with an improvement in symptoms after chemotherapy, compared with uncontrolled disease, the cause for the decline is not known from the available data.

Trial protocols versus clinical practice

In clinical trials, patients are randomised to different therapies, and the protocol encourages adherence to therapy to the limits of tolerance, though because of side-effects, many patients do not receive complete courses. However, in routine care, two factors may affect doses and costs. Firstly, clinicians will review continuation more critically, depending on objective response (e.g. radiological evidence of tumour shrinkage) and effect on symptoms, and treatment is probably stopped earlier than in trials in those patients who do not respond. Secondly, trials usually give mean or median benefits such as survival, and this may conceal the fact that only a minority of patients benefit (perhaps 20% have survival gains, and another 20% experience symptom relief).

These two factors will both tend to mean that, when continuation of treatment depends on clinical instinct and judgement as well as patient feedback, treatment may be stopped earlier in non-responders, making it more cost-effective. Observational studies in routine care would be needed to quantify this, but *Table 13* estimates the effect on cost-effectiveness if chemotherapy was stopped after 1.5 cycles in 60% of patients.

Chapter 5
Discussion

Statement of principal findings

The main findings of the rapid systematic review of DOC, GEM, PAX, VNB and other combined therapies for NSCLC are discussed in this section.

Docetaxel

DOC as second-line therapy, compared with BSC or VNB/IFOS, appears to be beneficial for patients with NSCLC when assessed using patient survival and QoL. Improvements in median survival were significantly higher for the patients receiving DOC, 75 mg/m^2, compared with BSC, but not for patients treated with DOC, 100 mg/m^2, or when compared with patients receiving VNB/IFOS. Similarly, 1- and 2-year survival was improved for patients receiving DOC, 75 mg/m^2 and 100 mg/m^2, compared with BSC and compared with VNB/IFOS. DOC provided improvements in all dimensions of QoL assessed, except for diarrhoea, with significant benefits in pain and fatigue scales. Haematological adverse effects from DOC were dose related and included anaemia, neutropenia, thrombocytopenia, infections and asthenia. Diarrhoea was the main non-haematological side-effect. Patients treated with DOC used significantly less opiate and non-opiate analgesics and tumour-related medications, other than for pain and palliative radiotherapy, but had higher rates of the use of anti-infective drugs. DOC appears reasonably cost-effective compared with BSC as second-line treatment, with the limited evidence favouring a 75-mg/m^2 dose over 3 cycles. Direct comparisons in the trial by Shepherd and co-workers[21] have indicated an incremental cost per LYS of £4234 compared with BSC. As first-line therapy, DOC (unlicensed) would appear to be expensive under a number of scenarios.

Gemcitabine

The benefits of GEM on patient survival and QoL for people with NSCLC varied depending on the comparator drug. While GEM alone did not lead to significant improvements in survival rates for patients when compared with CDDP+VP-16, it significantly increased survival time when compared with VP-16 alone. Similarly, studies of GEM+CDDP provided contradictory results. If the intervention was compared with CDDP alone, then there was a significant improvement in survival for GEM+CDDP, but when compared with the MITO+IFOS+CDDP arm, there was no significant improvement. The GEM+BSC intervention did not provide improved survival compared with BSC, which was unexpected, but GEM+BSC compared with BSC reduced the need for palliative radiotherapy. Haematological side-effects associated with GEM included anaemia, neutropenia and thrombocytopenia. Non-haematological side-effects included hair loss, infection, nausea and diarrhoea. There was no significant difference between the other interventions and GEM in terms of side-effects. GEM appears to offer reasonable cost-effectiveness compared with BSC under a range of scenarios in the cost-effectiveness analysis but does less well in the single-trial pairwise comparison. Although GEM+ CDDP dominated VNB+CDDP in the trial by Comella and co-workers,[50] it does less well in the range of assumptions in our cost-effectiveness model while still being reasonably cost-effective. However, GEM+CDDP is comparable with VNB+CDDP in terms of cost-effectiveness (relative to BSC), if it can be assumed that equal cycles are appropriate for each regimen and each VNB cycle consists of four administrations.

Paclitaxel

PAX appears to improve patient survival, whether median or 1-year survival, for people with NSCLC, when compared with VP-16+CDDP or BSC. Median survival of patients receiving PAX improved significantly when compared with BSC. In contrast, there was limited difference in median survival when PAX combinations were compared with MER, PIR or CDDP+VM-26. Improvements in 1- and 2-year survival were significant when compared with BSC. QoL appeared to be improved for patients receiving PAX, although improvements in functional ability were the only outcomes to differ significantly compared with BSC. Differences in adverse events depended on the interventions compared. Haematological side-effects of PAX included leucopenia, thrombocytopenia, neutropenia and anaemia, while non-haematological events from PAX included nausea/vomiting, diarrhoea, infection, cardiac events and myalgias. The PAX+CDDP regimens offer the largest gain in survival over BSC. Cost-effectiveness would tend to favour the

135-mg/m^2 dose, though this is based on data from a single trial.

Vinorelbine

VNB in combination with other drugs, particularly CDDP, was beneficial to patients with NSCLC when assessed on patient survival (median and 1-year survival). Limited difference was found when comparing patients based on QoL measures. Median survival was improved, although not significantly, for people receiving VNB+CDDP in combination or other VNB combinations. One-year survival was improved for people receiving VNB and VNB+CDDP, compared with FU+LV or CDDP alone or BSC. In contrast, a smaller proportion of people receiving VNB+CDDP+IFOS (compared with CDDP+MITO+VDS or CBDCA) and VNB+CBDCA (compared with VNB+CDDP) survived to 1 year. None of the differences were statistically significant.

Changes in QoL showed limited difference when comparing VNB with 5FU+LV and when comparing VNB+CDDP with EPI and CDDP. Comparison of VNB with BSC showed statistically significant improvements in QoL for patients on VNB for cognitive function, pain and medication, and dyspnoea, but significant worsening for constipation, nausea and vomiting, peripheral neuropathy and hair loss. Haematological and non-haematological adverse effects were more evident in people receiving VNB, including thrombocytopenia, leucopenia, anaemia, alopecia, nausea/vomiting and diarrhoea.

Combination therapies

The effectiveness of combination therapies in the treatment of people with NSCLC varied with the different drugs compared. CDDP+GEM+VNB was beneficial to patients when assessed on median patient survival and 1- and 2-year survival, compared with CDDP+EPI+VDS+LON, CDDP+GEM and CDDP+VNB. Differences in median survival were statistically significant in the comparison of CDDP+GEM+VNB and CDDP+EPI+VDS+LON. Although patients on VNB+IFOS had a higher median survival than those on PAX+IFOS, they had a lower 1-year survival rate than those on PAX+IFOS; neither differed significantly.

The only comparisons of QoL were for patients receiving CDDP+VNB+GEM compared with CDDP+EPI+VDS and patients receiving GEM+VNB compared with VNB; patients receiving CDDP+VNB+GEM and GEM+VNB were found to have improved QoL. In comparisons of combinations including CDDP+GEM with other combination therapies, CDDP+GEM combinations tended to have higher haematological and non-haematological toxicities than the other combination therapies. Haematological and non-haematological adverse effects were more common in patients treated with PAX+CBDA compared with patients treated with PAX+GEM. Similarly, VNB+IFOS was associated with more haematological and non-haematological adverse effects than PAX+IFOS.

VNB and VNB+CDDP offer reasonable incremental cost-effectiveness and retain this throughout our range of assumptions in the sensitivity analysis. In certain cases, they may even be cost-saving compared with BSC. VNB+CDDP offers additional survival compared to single-agent VNB, at arguably little additional cost. A small trial of VNB+CDDP with GEM+CDDP found in favour of GEM+CDDP but had a number of methodological factors that make interpretation difficult.

Strengths and limitations of the review

This review has certain strengths, including the following.

- It is independent of any vested interest.
- The review brings together the evidence on the effectiveness of four drugs for lung cancer and an economic evaluation, applying consistent methods of critical appraisal and presentation.
- The review was guided by the principles for undertaking a systematic review. Prior to undertaking the rapid review, the methods of the review were set out in a research protocol (appendix 1), which was commented on by an advisory group. The protocol defined the research question, inclusion criteria, quality criteria, data extraction process and methods employed to undertake the different stages of the review.
- An advisory group has informed the rapid review from its initiation, through the development of the research protocol and completion of the report.

In contrast, there were certain limitations placed upon the review.

- Due to differences in the design, duration, outcome measures and reporting of studies, synthesis of the included studies was through narrative review with no formal meta-analysis.

- Another restriction placed upon the review due to time constraints was the lack of follow-up with authors of studies to clarify information presented.
- The quality of the RCTs was assessed using the Jadad scale.[19] Although the Jadad scale includes key elements by which to assess the quality of RCTs, including randomisation, blinding and withdrawals/dropouts, it could be criticised for excluding other elements that may cause bias (e.g. not including the level of withdrawal/dropout). It has also been pointed out that the Jadad scale "gives more weight to the quality of reporting than to actual methodological quality".[82]

Other issues

Overall, approximately 47% of the critically appraised studies (16 of 34) reported QoL measures as either a primary or secondary end-point.[14,21,22,24,26,27,29,30,32,34,35,40,45,48,49,53] Although 13 of these studies evaluated QoL using standardised questionnaires, the standard of reporting could be regarded as high in only five of the studies.[14,27,30,32,48] This greatly reduces the number of trials in which QoL can be reliably considered for this review. It is imperative, for both quantifying the benefits and extrapolating cost-effectiveness of chemotherapy agents used in NSCLC, that trials adequately report changes in patients' QoL using standardised measurement instruments.

Are the new drugs effective in NSCLC?

If we based drug effectiveness on comparison with BSC, then the conclusion is that the drugs are effective. The evidence comes in part directly from the few trials of the drugs versus BSC, and in part indirectly from trials in which they were compared with existing regimens that are known to be better than BSC, notably the CDDP or combination studies.[13,83,84]

The effectiveness of all chemotherapy in NSCLC is modest. Cure is not possible yet, and the gains are in small extensions of survival and benefits in QoL, with drug treatment being less life-diminishing than uncontrolled growth of tumour or metastases.

Cost-effectiveness

The extension of life, though modest, is achieved at relatively low cost, and the cost per LYG would be well within the usual funding range for the NHS.

If an inverse weighting for duration of life was applied, then these drugs would become more cost-effective.

In routine clinical practice, in which the continuation of chemotherapy is decided on a continuing individual patient basis determined by tumour response, side-effects, clinical nous and patient choice, it is likely that chemotherapy would be stopped much earlier in non-responders than in the trials. This limits suffering due to side-effects and considerably improves the cost-effectiveness of care.

If oral versions of VNB and GEM become available, then their administration costs will fall, and cost-effectiveness will increase further. Once generic PAX appears, its cost is also expected to fall. If there was a shift in combination therapy from CDDP to CBDCA (which would require evidence of equivalence or improved cost-effectiveness), costs may fall via a reduction in inpatient costs.

Chapter 6
Conclusions

Implications for other parties

Implications for the NHS

Any increase in the proportion of patients with NSCLC who receive chemotherapy will happen at a time when their absolute numbers are slowly reducing, but because such a small proportion of patients are treated at present, any increase in that proportion will mean a considerable increase in absolute numbers treated. The costs of this increase will be in staff time, laboratory support and drugs. Although some regimens are cost-saving, the savings will be realised only if hospital beds are being closed (which in effect is a euphemism for reducing numbers of nurses in these wards, because nursing is the largest single component of bed costs). However, because bed numbers are already being reduced in many places, the main impact of chemotherapy may be to make it easier to reduce bed numbers (in effect achieving savings, but ones that were already planned; the funds may already have been reallocated and therefore would not be available for reinvestment in chemotherapy services).

More nursing time would be needed for counselling patients and obtaining informed consent, including for trials, and so there would be a switch of resources from one part of the service to another. In our economic model, we have included 1 hour of nurse time for informing and counselling patients before chemotherapy.

The NHS Cancer Plan[2] has already announced that the number of medical oncologists in England will increase by 141% (from 110 to 265) between 1999 and 2006, and that the number of nurses will also increase (no figure for cancer nurses given). The same document[2] states (on page 13, paragraph 36) that "the new funding announced in the NHS plan will mean that when NICE publishes its guidance on new cancer drugs in summer 2001, health authorities right across the country will be able to take full account of it". This implies that extra funding will be available.

Numbers to be treated

Without better information on current treatment patterns, it is difficult to estimate the total budgetary impact. Each year, there are about 26,500 new patients with NSCLC in England and Wales. If we assume that 10% are cured surgically, about 24,000 patients are left. Many will be unfit for chemotherapy, because of frailty due to age or co-morbidities. Others will decline chemotherapy. Some who start will stop because of side-effects. Many will not continue because of lack of response.

If 25% of those patients with non-resectable tumours were to be treated and if we start with a pessimistic assumption that there will be no realisable savings at all in real life, then the cost would be around £10 million per annum (based on VNB, 3 cycles per patient, using the marginal cost over BSC, but assuming that the reduced bed-days needed would be used for other purposes, with no monetary savings).

The variables that most affect the global impact on budgets are:

- the number of patients with inoperable NSCLC
- the proportion treated with chemotherapy
- treatment variables, including choice of drug, intensity of treatment, whether oral or intravenous, outpatient or inpatient, and frequency of serious side-effects requiring admission.

Hence, unless a much larger proportion of patients are treated, the extra cost is not great, even with our pessimistic assumptions about the reality of savings. There will be regional variations due to differences in smoking prevalence.

We sought advice from our expert advisers on what proportion of patients they would expect to be treated, taking into account not just the arrival of the new drugs but also practical barriers to implementation. Their estimates ranged from 10–15% to over 50%, but the differences were partly due to the denominators used, and partly from inclusion of chemotherapy given as adjuvant treatment with surgery or radiotherapy. Many patients are too old or frail for chemotherapy, and including them will drop the percentage to be treated. Others may not be referred to an oncologist. Our estimate of 25% of all patients with NSCLC is probably a bit too high for the next couple of years, in most places, but might be a bit too low in 3–5 years' time.

Factors that would affect implementation

If NICE were to support greater use of chemotherapy with the newer drugs, the number of patients treated would be influenced by the following factors.

1. Referral for chemotherapy

There still appears to be a common perception among many GPs and physicians that chemotherapy is toxic and ineffective in lung cancer, and this perception will affect referral to oncology services. It could be argued that this perception is valid – because none of the current drugs offer cure and because survival is increased by only a few months. However, this perception would likely be changed by a clear statement from NICE that chemotherapy is clinically effective, in terms of modest but proportionally important gains in survival, accompanied by gains in QoL, and also inexpensive compared with many other NHS activities. The influence of medical and clinical oncologists, and perhaps particularly of chest physicians with a special interest in lung cancer, is also reducing the previous pessimism.

2. Patient expectations

Advice from the Roy Castle Lung Cancer Foundation suggests that there may be a general belief among patients that the side-effects of treatment outweigh the benefits. This is in accord with previous reports that most patients who were fully informed about the benefits and side-effects of the older drugs declined to enter trials.[85] However, this may be changing, partly as information about the newer drugs spreads and partly as the availability of information to patients increases, in particular from the Internet. A recent report from the USA noted that 22% of patients with NSCLC received chemotherapy.[86] Again, guidance from NICE would be likely to change perceptions.

3. Shortages of trained staff or staff time

This shortage of staff refers to oncologists, chest physicians, nurses and pharmacists. Most respiratory physicians would refer patients to oncologists for chemotherapy, but time is required for unhurried discussion with patients, in order to provide fully informed consent. Nurses may have a key role here, and we are aware of the spread of specialist nurses, who may not only counsel patients but administer the chemotherapy.

The increase in the number of oncologists has been referred to above (see *Implications for the NHS*).

Skilled pharmacy support is needed, not least for preparation of drug infusions. Because of concern about possible harm to staff that might arise from chronic low-grade exposure to some of the anticancer drugs, the reconstitution is now done under controlled conditions. There is competition to recruit pharmacists, for example, from primary care groups.

4. Beds

Several members of our advisory panel noted that shortage of beds for inpatient or day-case chemotherapy was a constraint. With the older drugs such as CDDP combinations, patients have to be admitted for fluid infusion so that they are sufficiently well hydrated to reduce side-effects. The newer drugs can all be delivered on an outpatient basis, either as single agents or in combination. This provides an opportunity for outpatient chemotherapy to ease the bed pressure and allow more patients to be treated more promptly. The cost implications are twofold – an increase in numbers treated, but a decrease in inpatient bed-days, partly from reduced inpatient chemotherapy and partly from reduced bed-days needed compared with BSC. Unpublished data from the UK Gemcitabine Study shows that, compared with BSC, those patients who had chemotherapy needed much less palliative radiotherapy than the BSC arm: 42% of the BSC group had radiotherapy, compared with 7% of the GEM group. The number of bed-days used was less in the GEM arm than in the BSC one. Hence, one impact of chemotherapy may be to reduce radiotherapy workload and waiting times (Thatcher N, Christie Hospital, Manchester: personal communication, 2000).

5. Budgets

Even if the total cost to the NHS is modest, the impact on some local budgets, such as pharmacy, will be considerable. If sufficient new funds are not available, an increase in chemotherapy would need to be funded by reallocation of resources – always difficult in practice.

Factors relevant to NHS policy

The NHS Cancer Plan[2] has been referred to already. Section 6.6 of the plan refers to the forthcoming advice from NICE and expresses a desire to end "the postcode prescribing lottery". Paragraph 6.27 notes that a national minimum data set for lung cancer will be introduced in 2001, and these data will presumably reveal any geographical variations in treatment and outcomes.

The National electronic Library for Health (NeLH) plans to provide reliable information for patients on the Internet, starting with cancer. The first five cancers to be covered will include lung cancer.

Equity issues have been referred to above and in particular differences in survival. Another issue that may emerge is the age range of patients treated. New and more gentle forms of chemotherapy may allow more elderly patients to be treated, and evidence from Italy (ELVIS) shows benefit in this group.[53] It should be noted that 40% of people with lung cancer are over 75 years of age (*Table 1*).

Recommendations for research

The need for trials

The more patients who are recruited to trials, the faster results will emerge. Research will be a slow process involving repeated trials, sometimes of a new drug, sometimes of a combination or of alternating drugs, sometimes of different methods of assessment or staging, using different investigations, different patient groups, different methods of assessing, for example, QoL. Hence, there is a good case for encouraging more patients to enter trials and for experimental treatments to be provided only through trials. Lung cancer has proved to be more resistant to chemotherapy than some other cancers, such as childhood leukaemia or testicular cancer. Getting better results may come slowly, in small but cumulatively useful increments. There will be retreats as well as advances. The infrastructure for research needs to be considered, but is beyond the remit of this review. Trials are underway, including the UK-based 'Big Lung Trial'.

Specific research areas

A number of research needs have been identified by this review, including the need for:

1. good-quality RCTs of drugs or combinations thereof among different subgroups of patients, with QoL measured using a validated method
2. studies of the use of chemotherapy in combination with radiotherapy
3. methods for assessing the methodological quality of QoL studies
4. prospective cost collection and economic analysis as part of future studies of treatment.

Who decides?

At present, it seems that doctors, presumably largely respiratory physicians, are the main movers in determining the frequency of chemotherapy in NSCLC. Two studies have looked at patient views. Slevin and co-workers[87] compared the views of doctors, nurses and patients, and found that patients would choose treatment at a lower level of benefit than clinicians:

> "Faced with the reality of the diagnosis...patients are likely to accept any treatment that offers them some possible benefit and hope, however slight."

There can be at least two explanations for the differences. Firstly, the level of knowledge of side-effects – the patients had cancer but had not yet had chemotherapy. Secondly, the clinicians may have taken cost and cost-effectiveness into account.

Silvestri and colleagues[77] interviewed patients who had had at least 1 cycle of platinum-based regimens, and found a wide range of thresholds for accepting chemotherapy. Some patients would not accept chemotherapy if it gave only 3 months of survival benefit, but they would if it improved QoL. But a few patients would accept chemotherapy with severe toxicity even if it gave only short gains in duration of life, and two-thirds would accept chemotherapy (and it is worth noting that this was the older more toxic forms rather than the drugs reviewed here) if it relieved symptoms with no extension of life.

It is likely that the advent of gentler forms of chemotherapy will increase the number of patients who choose to have it.

Acknowledgements

We are very grateful to the advisory panel below, who provided expert advice and comments on the draft of this report, but we absolve them from any shortcomings in the final report, which remains the responsibility of the Southampton Health Technology Assessments Centre alone. Dr Norman Waugh is guarantor.

We thank the Royal College of Physicians of London, the Cancer Research Campaign (CRC), and the Faculty of Clinical Oncology of the Royal College of Radiologists for nominating expert advisers:

Dr J Baird
Director of Patient Care
Roy Castle Lung Cancer Foundation
Glasgow

Dr Frances Bowen
Consultant in Respiratory and General Medicine
Hammersmith Hospital
London

Dr Doug Coyle
Principal Investigator, Health Economics
Clinical Epidemiology Unit
The Ottawa Hospital
Ottawa, Canada

Dr Michael Cullen
Consultant and Hon. Reader in Medical Oncology
Cancer Centre at the Queen Elizabeth Hospital
Birmingham

Dr Anna Gregor
Consultant in Clinical Oncology
Western General Hospital
Edinburgh

Dr P Hopwood
Hon. Consultant Psychiatrist
CRC Psychological Medicine Group
Christie Hospital
Manchester

Dr Fergus MacBeth
Consultant Oncologist
Velindre Hospital
Cardiff

Ms Kirsten Major
Health Economist
Ayr and Arran Health Board

Dr Ben Marshall
Consultant in Respiratory Medicine
Southampton General Hospital
Southampton

Dr Martin Muers
Consultant in Respiratory Medicine
Leeds

Professor IE Smith
Professor of Cancer Medicine
The Royal Marsden Hospital
Sutton

Professor N Thatcher
Professor of Oncology
CRC Department of Medical Oncology
Christie CRC Research Centre
Christie Hospital
Manchester.

A number of the above advisers declared possible competing interests in advance, such as involvement in trials sponsored by industry.

Useful comments were obtained from members of the NICE Appraisal Committee and Secretariat at the meeting in February 2001, and from the consultees after the provisional appraisal decision and the previous version of this review had been circulated for comment.

We are grateful to David Carson, Rob Hudson, Marlyn Foster and Lynda Dick of the former SHPIC costing unit, and to Dr John Winter and colleagues at King's Cross Hospital, Dundee, for cost data, as well as to the Southeast Scotland Lung Cancer Group, and in particular Dr Ann Gould and Dr Ann Cull, for access to unpublished resource utilisation data.

We would also like to thank the following people for information support: Dr Pam Royle and Ms Liz Hodson, Information Service, Wessex Institute for Health Research and Development.

Thanks also to Dr Emma Loveman and Linda Durrant, Wessex Institute for Health Research and Development, for help in preparing the draft report.

We are also indebted to the referees for their comments on the report. The views expressed in this report are those of the authors, who are also responsible for any errors.

Authorship

The report's authorship is as follows:

- protocol: N Waugh, A Clegg, P Hewitson, M Sidhu and DA Scott
- searching: P Hewitson and DA Scott
- inclusion criteria: P Hewitson, M Sidhu, DA Scott and A Clegg
- data extraction: M Sidhu, A Clegg, P Hewitson and DA Scott
- economic analysis: D Scott and M Sidhu
- first draft: DA Scott, A Clegg, M Sidhu, P Hewitson and N Waugh
- final draft: N Waugh, D Scott and A Clegg.

References

1. Coleman MP, Babb P, Domiecki P, Grosclaude P, Honjo S, Jones J, *et al.* Cancer survival trends in England and Wales, 1971–1995: deprivation and NHS region. London: The Stationery Office; 1999.
2. Department of Health. The NHS Cancer Plan. A plan for investment. A plan for reform. London: The Stationery Office; 2000.
3. Office for National Statistics. Health statistics quarterly. London: The Stationery Office; 2000.
4. Department of Health. NHS performance indicators. 2001. URL: http://www.doh.gov.uk/nhsperformanceindicators/
5. Acheson D (chairman). Independent inquiry into inequalities in health: report. London: The Stationery Office; 1998.
6. Welsh Thoracic Society. Lung cancer in Wales: a longitudinal retrospective audit of the hospital management of patients newly diagnosed with lung cancer in Wales in 1996. Cardiff: Clinical Effectiveness Support Unit (CESU); 2000.
7. Crook A, Duffy A, Girling DJ, Souhami RL, Parmar MK. Survey on the treatment of non-small cell lung cancer (NSCLC) in England and Wales. *Eur Respir J* 1997;**10**:1552–8.
8. Souquet PJ, Chauvin F, Boissel JP, Cellerino R, Cormier Y, Ganz PA, *et al.* Polychemotherapy in advanced non small cell lung cancer: a meta-analysis. *Lancet* 1993;**342**:19–21.
9. Smith IE. Palliative chemotherapy for advanced non-small cell lung cancer. *BMJ* 1994;**308**:429–30.
10. Hardy JR, NobleT, Smith IE. Symptom relief with moderate dose chemotherapy (mitomycin-C, vinblastine and cisplatin) in advanced non small cell lung cancer. *Br J Cancer* 1989;**60**:764–6.
11. Marino P, Pampallona S, Preatoni A, Cantoni A, Invernizzi F. Chemotherapy vs supportive care in advanced non-small-cell lung cancer. Results of a meta-analysis of the literature. *Chest* 1994;**106**:861–5.
12. Thatcher N, Anderson H, Betticher DC, Ranson M. Symptomatic benefit from gemcitabine and other chemotherapy in advanced non-small cell lung cancer: changes in performance status and tumour-related symptoms. *Anticancer Drugs* 1995;**6**(Suppl 6):39–48.
13. Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patient data from 52 randomised clinical trials. *BMJ* 1995;**311**:899–909.
14. Anderson H, Hopwood P, Stephens RJ, Thatcher N, Cottier B, Nicholson M, *et al.* Gemcitabine plus best supportive care (BSC) vs BSC in inoperable non-small cell lung cancer – a randomized trial with quality of life as the primary outcome. *Br J Cancer* 2000;**83**:447–53.
15. Jaakimainen L, Goodwin PJ, Pater J, Warde P, Murray N, Rapp E. Counting the costs of chemotherapy in a National Cancer Institute of Canada randomised trial in NSCLC. *J Clin Oncol* 1990;**8**:1301–9.
16. Macbeth F. Managing patients with lung cancer. All evidence was considered when COIN guidelines were drawn up. *BMJ* 2000;**320**:1604.
17. O'Brien MER, Cullen M. Managing patients with lung cancer. Guidelines must help bring us in line with European standards. *BMJ* 2000;**320**:1604.
18. The National Cancer Guidance Group. Improving outcomes in lung cancer: the research evidence. London: NHS Executive, Department of Health; 1998.
19. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
20. NHS Centre for Reviews and Dissemination. 2000. URL: http://agatha.york.ac.uk/darehp.htm
21. Shepherd FA, Dancey J, Ramlau R, Mattson K, Gralla R, O'Rourke M, *et al.* Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. *J Clin Oncol* 2000;**18**:2095–103.
22. Roszkowski K, Pluzanska A, Krzakowski M, Smith AP, Saigi E, Aasebo U, *et al.* A multicenter, randomized, phase III study of docetaxel plus best supportive care versus best supportive care in chemotherapy-naive patients with metastatic or non-resectable localized non-small cell lung cancer (NSCLC). *Lung Cancer* 2000;**27**:145–57.
23. Fossella FV, Devore R, Kerr RN, Crawford J, Natale M, Dunphy F, *et al.* Randomized phase III trial of docetaxel versus vinorelbine or ifosfamide in patients with advanced non-small-cell lung cancer previously treated with platinum-containing chemotherapy regimes. *J Clin Oncol* 2000;**18**:2354–62.

24. Bokkel-Huinink WW, Bergman B, Chemaissani A, Dornoff W, Drings P, Kellokumpu-Lehtinen PL, *et al.* Single-agent gemcitabine: an active and better tolerated alternative to standard cisplatin-based chemotherapy in locally advanced or metastatic non-small cell lung cancer. *Lung Cancer* 1999;**26**:85–94.

25. Manegold C, Bergman B, Chemaissani A, Dornoff W, Drings P, Kellokumpu LP, *et al.* Single-agent gemcitabine versus cisplatin–etoposide: early results of a randomised phase II study in locally advanced or metastatic non-small-cell lung cancer. *Ann Oncol* 1997;**8**:525–9.

26. Cardenal F, Lopez-Cabrerizo MP, Anton A, Alberola V, Massuti B, Carrato A, *et al.* Randomized phase III study of gemcitabine–cisplatin versus etoposide– cisplatin in the treatment of locally advanced or metastatic non-small-cell lung cancer. *J Clin Oncol* 1999;**17**:12–18.

27. Crino L, Scagliotti GV, Ricci S, De Marinis F, Rinaldi M, Gridelli C, *et al.* Gemcitabine and cisplatin versus mitomycin, ifosfamide, and cisplatin in advanced non-small-cell lung cancer: a randomized phase III study of the Italian Lung Cancer Project. *J Clin Oncol* 1999;**17**:3522–30.

28. Perng RP, Chen YM, Ming LJ, Tsai CM, Lin WC, Yang KY, *et al.* Gemcitabine versus the combination of cisplatin and etoposide in patients with inoperable non-small-cell lung cancer in a phase II randomized study. *J Clin Oncol* 1997;**15**:2097–102.

29. Sandler AB, Nemunaitis J, Denham C, von Pawel J, Cormier Y, Gatzemeier U, *et al.* Phase III trial of gemcitabine plus cisplatin versus cisplatin alone in patients with locally advanced or metastatic non-small-cell lung cancer. *J Clin Oncol* 2000;**18**:122–30.

30. Bonomi P, Kim K, Fairclough D, Cella D, Kugler J, Rowinsky E, *et al.* Comparison of survival and quality of life in advanced non-small-cell lung cancer patients treated with two dose levels of paclitaxel combined with cisplatin versus etoposide with cisplatin: results of an Eastern Cooperative Oncology Group trial. *J Clin Oncol* 2000;**18**:623–31.

31. Chang AY, Kim K, Glick J, Anderson T, Karp D, Johnson D. Phase II study of taxol, merbarone, and piroxantrone in stage IV non-small-cell lung cancer: the Eastern Cooperative Oncology Group results. *J Natl Cancer Inst* 1993;**85**:388–94.

32. Ranson M, Davidson N, Nicolson M, Falk S, Carmichael J, Lopez P, *et al.* Randomized trial of paclitaxel plus supportive care versus supportive care for patients with advanced non-small cell lung cancer. *J Natl Cancer Inst* 2000;**92**:1074–80.

33. Postmus PE, Giaccone G, Debruyne C, Sahmoud T, Splinter TA, van Zandwijk N. Results of the phase II EORTC study comparing paclitaxel/cisplatin with teniposide/cisplatin in patients with non-small cell lung cancer. EORTC Lung Cancer Cooperative Group. *Semin Oncol* 1996;**23**(Suppl 12):10–13.

34. Gatzemeier U, von Pawel J, Gottfried M, ten Velde GPM, Mattson K, DeMarinis F, *et al.* Phase III comparative study of high-dose cisplatin versus a combination of paclitaxel and cisplatin in patients with advanced non-small-cell lung cancer. *J Clin Oncol* 2000;**18**:3390–9.

35. Giaccone G, Splinter TA, Debruyne C, Kho GS, Lianes P, van Zandwijk N, *et al.* Randomized study of paclitaxel–cisplatin versus cisplatin–teniposide in patients with advanced non-small-cell lung cancer. The European Organisation for Research and Treatment of Cancer Lung Cancer Cooperative Group. *J Clin Oncol* 1998;**16**:2133–41.

36. Baldini E, Tibaldi C, Ardizzoni A, Salvati F, Antilli A, Portalone L, *et al.* Cisplatin–vindesine–mitomycin (MVP) vs cisplatin–ifosfamide–vinorelbine (PIN) vs carboplatin–vinorelbine (CaN) in patients with advanced non-small-cell lung cancer (NSCLC): a FONICAP randomized phase II study. Italian Lung Cancer Task Force (FONICAP). *Br J Cancer* 1998;**77**:2367–70.

37. Colleoni M, Vicario G, Pancheri F, Sgarbossa G, Nelli P, Manente P. A randomized phase II trial of cisplatinum plus mitomycin-C plus vinorelbine and carboplatin plus vinorelbine in advanced non-small cell lung cancer. *Int J Oncol* 1997;**10**:619–22.

38. Colucci G, Gebbia V, Galetta D, Riccardi F, Cariello S, Gebbia N. Cisplatin and vinorelbine followed by ifosfamide plus epirubicin vs the opposite sequence in advanced unresectable stage III and metastatic stage IV non-small-cell lung cancer: a prospective randomized study of the Southern Italy Oncology Group (GOIM). *Br J Cancer* 1997;**76**:1509–17.

39. Comella P, Frasci G, De Cataldis G, Panza N, Cioffi R, Curcio C, *et al.* Cisplatin/carboplatin + etoposide + vinorelbine in advanced non-small-cell lung cancer: a multicentre randomised trial. Gruppo Oncologico Campano. *Br J Cancer* 1996;**74**:1805–11.

40. Crawford J, O'Rourke M, Schiller JH, Spiridonidis CH, Yanovich S, Ozer H, *et al.* Randomized trial of vinorelbine compared with fluorouracil plus leucovorin in patients with stage IV non-small-cell lung cancer. *J Clin Oncol* 1996;**14**:2774–84.

41. Depierre A, Chastang C, Quoix E, Lebeau B, Blanchon F, Paillot N, *et al.* Vinorelbine versus vinorelbine plus cisplatin in advanced non-small cell lung cancer: a randomized trial. *Ann Oncol* 1994;**5**:37–42.

42. Furuse K, Fukuoka M, Kuba M, Yamori S, Nakai Y, Negoro S, *et al.* Randomized study of vinorelbine (VRB) versus vindesine (VDS) in previously untreated stage IIIB or IV non-small-cell lung cancer (NSCLC). The Japan Vinorelbine Lung Cancer Cooperative Study Group. *Ann Oncol* 1996;**7**:815–20.

43. Le Chevalier T, Brisgand D, Douillard JY, Pujol JL, Alberola V, Monnier A, *et al.* Randomized study of vinorelbine and cisplatin versus vindesine and cisplatin versus vinorelbine alone in advanced non-small-cell lung cancer: results of a European multicenter trial including 612 patients. *J Clin Oncol* 1994;**12**:360–7.

44. Lorusso V, Pezzella G, Catino AM, Guida M, Scoditti S, Lorenzo R, *et al.* Results of a clinical multicentric randomized phase II study of non-small cell lung cancer treated with vinorelbine–cisplatin versus vinorelbine alone. *Int J Oncol* 1995;**6**:65–8.

45. Martoni A, Guaraldi M, Piana E, Strocchi E, Petralia A, Busutti L, *et al.* Multicenter randomized clinical trial on high-dose epirubicin plus cis-platinum versus vinorelbine plus cis-platinum in advanced non small cell lung cancer. *Lung Cancer* 1998;**22**:31–8.

46. Pérol M, Guerin JC, Thomas P, Poirier R, Carles P, Robinet G, *et al.* Multicenter randomized trial comparing cisplatin–mitomycin–vinorelbine versus cisplatin–mitomycin–vindesine in advanced non-small cell lung cancer. 'Groupe Français de Pneumo-Cancerologie'. *Lung Cancer* 1996;**14**:119–34.

47. Wozniak AJ, Crowley JJ, Balcerzak SP, Weiss GR, Spiridonidis CH, Baker LH, *et al.* Randomized trial comparing cisplatin with cisplatin plus vinorelbine in the treatment of advanced non-small-cell lung cancer: a Southwest Oncology Group study. *J Clin Oncol* 1998;**16**:2459–65.

48. The Elderly Lung Cancer Vinorelbine Italian Study Group (ELVIS). Effects of vinorelbine on quality of life and survival of elderly patients with advanced non-small-cell lung cancer. *J Natl Cancer Inst* 1999;**91**:66–72.

49. Comella P, Frasci G, Panza N, Manzione L, Lorusso V, Di Rienzo G, *et al.* Cisplatin, gemcitabine, and vinorelbine combination therapy in advanced non-small-cell lung cancer: a phase II randomized study of the Southern Italy Cooperative Oncology Group. *J Clin Oncol* 1999;**17**:1526–34.

50. Comella P, Frasci G, Panza N, Manzione L, De Cataldis G, Cioffi R, *et al.* Randomized trial comparing cisplatin, gemcitabine, and vinorelbine with either cisplatin and gemcitabine or cisplatin and vinorelbine in advanced non-small-cell lung cancer: interim analysis of a phase III trial of the Southern Italy Cooperative Oncology Group. *J Clin Oncol* 2000;**18**:1451–7.

51. Kosmidis P, Mylonakis N, Dimopoulos A, Pavlidis N, Fountzilas G, Samantas E, *et al.* Combination chemotherapy with paclitaxel plus carboplatin versus paclitaxel plus gemcitabine in inoperable non-small cell lung cancer: a phase III randomized study. Preliminary results. Hellenic Cooperative Oncology Group. *Semin Oncol* 2000;**27**(Suppl 2):3–8.

52. Perry MC, Ihde DC, Herndon JE, Grossbard ML, Grethein SJ, Atkins JN, *et al.* Paclitaxel/ifosfamide or navelbine/ifosfamide chemotherapy for advanced non-small cell lung cancer: CALGB 9532. *Lung Cancer* 2000;**28**:63–8.

53. Frasci G, Lorusso V, Panza N, Comella P, Nicolella G, Bianco A, *et al.* Gemcitabine plus vinorelbine versus vinorelbine alone in elderly patients with advanced non-small-cell lung cancer. *J Clin Oncol* 2000;**18**:2529–36.

54. Berthelot JM, Will BP, Evans WK, Coyle D, Earle CC, Bordeleau L. Decision framework for chemotherapeutic interventions for metastatic non-small-cell lung cancer. *J Natl Cancer Inst* 2000;**92**:1321–9.

55. Evans WK. Cost-effectiveness of gemcitabine in stage IV non-small cell lung cancer: an estimate using the Population Health Model lung cancer module. *Semin Oncol* 1997;**24**(Suppl 7):S7.

56. Copley-Merriman C, Corral J, King K, Whiteside R, Voi M, Dorr FA, *et al.* Economic value of gemcitabine compared to cisplatin and etoposide in non-small cell lung cancer. *Lung Cancer* 1996;**14**:45–61.

57. Copley-Merriman C, Martin C, Johnson D, Sacristan JA, Drings P, Bosanquet N. Economic value of gemcitabine in non-small cell lung cancer. *Semin Oncol* 1996;**5**:90–8.

58. Tennvall GR, Fernberg JO. Economic evaluation of gemcitabine single agent therapy compared with standard treatment in stage IIIB and IV non-small cell lung cancer. *Med Oncol* 1998;**15**:129–36.

59. Palmer AJ, Brandt A. The cost-effectiveness of four cisplatin-containing chemotherapy regimens in the treatment of stages IIIB and IV non-small cell lung cancer: an Italian perspective. *Monaldi Arch Chest Dis* 1996;**51**:279–88.

60. Sacristan JA, Kennedy MT, Rosell R, Cardenal F, Anton A, Lomas M, *et al.* Economic evaluation in a randomized phase III clinical trial comparing gemcitabine/cisplatin and etoposide/cisplatin in non-small cell lung cancer. *Lung Cancer* 2000;**28**:97–107.

61. Evans WK. An estimate of the cost effectiveness of gemcitabine in stage IV non-small cell lung cancer. *Semin Oncol* 1996;**23**(Suppl 10):82–9.

62. Smith TJ, Hillner BE, Neighbors DM, McSorley PA, Le Chevalier T. Economic evaluation of a randomized clinical trial comparing vinorelbine, vinorelbine plus cisplatin, and vindesine plus cisplatin for non-small-cell lung cancer. *J Clin Oncol* 1995;**13**:2166–73.

63. Hillner BE, Smith TJ. Cost-effectiveness analysis of three regimens using vinorelbine (navelbine) for non-small cell lung cancer. *Semin Oncol* 1996;**23**(Suppl 5):25–30.

64. Evans WK, Le Chevalier T. The cost-effectiveness of navelbine alone or in combination with cisplatin in comparison to other chemotherapy regimens and best supportive care in stage IV non-small cell lung cancer. *Eur J Cancer* 1996;**32A**:2249–55.

65. Evans WK. A comparison with standard chemotherapy and best supportive care: cost-effectiveness of vinorelbine alone or vinorelbine plus cisplatin for stage IV NSCLC. *Oncology* 1998;**12**(Suppl 5): 18–26.

66. Vergnenegre A, Pérol M, Pham E. Cost analysis of hospital treatment – two chemotherapic regimens for non-surgical non-small cell lung cancer. GFPC (Groupe Français Pneumo Cancerologie). *Lung Cancer* 1996;**14**:31–44.

67. Earle CC, Evans WK. Cost-effectiveness of paclitaxel plus cisplatin in advanced non-small-cell lung cancer. *Br J Cancer* 1999;**80**:815–20.

68. Annemans L, Giaccone G, Vergnenegre A. The cost-effectiveness of paclitaxel + cisplatin is similar to that of teniposide + cisplatin in advanced non-small cell lung cancer: a multicountry analysis. *Anticancer Drugs* 1999;**10**:605–15.

69. Berrino F. Survival of cancer patients in Europe: the Eurocare study. In: Berrino F, Sant M, Verdecchia A, Copocaccia R, Hakulinen T, Esteve J, editors. Survival of cancer patients in Europe. Lyon: International Agency for Research on Cancer (IARC) Scientific Publications; 1995. Report No.: 132.

70. Netten A, Curtis L. Unit costs of health and social care. Canterbury: Personal Social Services Research Unit, University of Kent; 2000.

71. Dinnes J, Waugh N, Hudson R, Wordsworth S. Breast cancer. Aberdeen: Scottish Health Purchasing Information Centre; 1997. URL: http://www.nhsconfed.net/shpic

72. Evans WK. Cost-effectiveness of vinorelbine alone or vinorelbine plus cisplatin for stage IV NSCLC. *Oncology (Huntingt)* 1998;**12**(Suppl 4):18–25.

73. Wolstenholme JL,Whynes DK. The hospital costs of treating lung cancer in the United Kingdom. *Br J Cancer* 1999;**80**:215–18.

74. Le Chevalier T, Pujol JL, Douillard JY, Alberola V, Monnier A, Riviere A, *et al.* A three-arm trial of vinorelbine plus cisplatin, vindesine plus cisplatin, and single-agent vinorelbine in the treatment of non-small cell lung cancer: an expanded analysis. *Semin Oncol* 1994;**21**(Suppl 10):28–33.

75. Manegold C, Drings P, von Pawel J, Ricci S, Dornoff W, van Walree N, *et al.* A randomized study of gemcitabine monotherapy versus etoposide/cisplatin in the treatment of locally advanced or metastatic non-small cell lung cancer. *Semin Oncol* 1997;**24**(Suppl 8):S8.

76. Gore SM, Fontaine O, Pierce NF. Impact of rice based oral rehydration solution on stool output and duration of diarrhoea: meta-analysis of 13 clinical trials. *BMJ* 1992;**304**:287–91.

77. Silvestri G, Pritchard R, Welch HG. Preferences for chemotherapy in patients with advanced non-small cell lung cancer. *BMJ* 1998;**317**:771–5.

78. Raby B, Pater J, Mackillop W. Does knowledge change practice? Another look at the management of non-small cell lung cancer. *J Clin Oncol* 1995;**13**:1904–11.

79. Rosenthal MA, Webster VJ, Gebski RC, Stuart-Harris AO, Langlands AO, Boyages J. The cost of treating small cell lung cancer. *Med J Aust* 1992;**156**:605–10.

80. Waugh N, Scott D. How should different life expectancies be valued? *BMJ* 1998;**316**:1316.

81. Brouwer W, Van Hout B. Diminishing marginal utility of health effects. *BMJ* 1998;**317**:1155.

82. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;**282**:1054–60.

83. Thatcher N, Hopwood P, Anderson H. Improving quality of life in patients with non-small cell lung cancer: research experience with gemcitabine. *Eur J Cancer* 1997;**33**(Suppl 1):S8–13.

84. van Zandwijk N, Giaccone G. Treatment of metastatic non-small cell lung cancer. *Curr Opin Oncol* 1996;**8**:120–5.

85. Spiro SG, Gower NH, Evans MT, Facchini FM, Rudd RM. Recruitment of patients with lung cancer into a randomised clinical trial: experience at two centres. *Thorax* 2000;**55**:463–5.

86. Earle CC, Venditti LN, Neumann PJ, Gelber RD, Weinstein MC, Potosky AL, *et al.* Who gets chemotherapy for metastatic lung cancer? *Chest* 2000;**117**:1239–46.

87. Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ, *et al.* Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public. *BMJ* 1990;**300**:1458–60.

88. Koch P, Johnson N, Van Schaik J, Andersen S, Blatter J, Bosanquet N, *et al.* Gemcitabine: clinical and economic impact in inoperable non-small cell lung cancer. *Anticancer Drugs* 1995; **6**(Suppl 6):49–54.

89. Earle CC, Evans WK. A comparison of the costs of paclitaxel and best supportive care in stage IV non-small-cell lung cancer. *Cancer Prev Control* 1997;**1**:282–8.

90. Gridelli C, Perrone F, Gallo C, Rossi A, Barletta E, Barzelloni ML, *et al.* Single-agent gemcitabine as second-line treatment in patients with advanced non small cell lung cancer (NSCLC): a phase II trial. *Anticancer Res* 1999;**19**:4535–8.

Studies reported as abstracts or conference presentations

Anderson H, Cottier B, Nicholson M, Milroy R, Maughan T, Bond M, *et al.* Phase III study of gemcitabine (Gemzar) versus best supportive care (BSC) in advanced non-small cell lung cancer (NSCLC). *Lung Cancer* 1997;**18**:9.

Berghmans T, Ninane V. Gemcitabine combined with cisplatin and carboplatin (CCG) or ifosfamide (IG): a pilot study preparing a randomised phase III ELCWP study in stage IV non small-cell lung carcinoma (NSCLC). *Eur Respir J* 1998;**12**(Suppl 28):251S–2S.

Bilaceroglu S, Kose T. Cisplatin/etoposide versus cisplatin/vinorelbine versus cisplatin/paclitaxel in advanced non-small cell lung cancer: a phase III randomized trial. *Eur Respir J* 1998;**12**(Suppl 28):8S.

Clerici M, Frontini L, Ianniello GP, Maiorino L, Piantedosi FV, Cigolari S, *et al.* Vinorelbine plus best supportive care (BSC) vs BSC in the treatment of advanced non small cell lung cancer (NSCLC) elderly patients. Results of a phase III randomized trial. *Ann Oncol* 1998;**9**(Suppl 3):62.

Comella G, Frasci G, Panza N, Nicolella G, Manzione L, De Cataldis G, *et al.* Cisplatin–gemcitabine vs cisplatin–gemcitabine–vinorelbine vs cisplatin–gemcitabine in advanced non-small-cell lung cancer. Final analysis of a Southern Italy Cooperative Oncology Group (SICOG) phase III trial. *Ann Oncol* 2000;**11**(Suppl 4):107.

Comella P, Panza N, Nicolella GP, Manzione L, Lorusso V, Cioffi R, *et al.* Cisplatin–gemcitabine–vinorelbine (PGV) vs cisplatin–vinorelbine (PV) vs cisplatin–gemcitabine (PG) in advanced non-small cell lung cancer (NSCLC). Interim analysis of a SICOG phase III trial. *Eur J Cancer* 1999;**35**(Suppl 4):S247.

Costa A, Barradas P, Cristovao M, Alves P, Melo MJ. Preliminary results of a randomised phase III trial comparing four cisplatin (p)-based regimens in the treatment of locally advanced and metastatic non-small cell lung cancer (NSCLC). *Lung Cancer* 2000;**29**:27.

Douillard JY. A randomized multicenter European study of paclitaxel with either cisplatin or carboplatin in advanced or metastatic non-small cell lung cancer: initial presentation. *Semin Oncol* 1999;**26**(Suppl 2):76.

Frasci G, Lorusso V, Panza N, Comella P, De Cataldis G, Micillo E, *et al.* Gemcitabine plus vinorelbine (GV) yields better survival than vinorelbine (V) alone in elderly non-small cell lung cancer (NSCLC) patients. Final analysis of a Southern Italy Cooperative Group (SICOG) phase III trial. *Ann Oncol* 2000;**11**(Suppl 4):108.

Frasci G, Lorusso V, Panza N, Decataldis G, Del Gaizo F, Nicolella GP, *et al.* Gemcitabine plus vinorelbine (GV) vs vinorelbine (V) alone in elderly or frail non-small cell lung cancer (NSCLC) patients. Interim analysis of a SICOG phase III trial. *Eur J Cancer* 1999; **35**(Suppl 4):S249.

Furuse K, Kawahara M, Nishiwaki Y, Horai T, Saijo N, Hasegawa K. Randomized study of vinorelbine or vindesine with cisplatin and mitomycin as induction chemotherapy in NSCLC. *Br J Cancer* 1998;**78**(Suppl 2):5.

Gatzemeier U, Rosell R, Betticher D, Keppler U, Macha HN, Pirker R, *et al.* Randomized pan-European trial comparing paclitaxel (TAX)/carboplatin (CAR) versus paclitaxel/cisplatin (CIS) in advanced non-small cell lung cancer (NSCLC). *Eur J Cancer* 1999; **35**(Suppl 4):S246.

Gridelli C, Ianniello GP, Maiorino L, Brancaccio L, Cigolari S, Bilancia D, *et al.* Vinorelbine (VNR) plus best supportive care (BSC) vs BSC in the treatment of advanced non-small cell lung cancer (NSCLC) elderly patients (Pts). Results of a phase III randomized trial. *Br J Cancer* 1998;**78**(Suppl 2):3.

Gridelli C, Ianniello GP, Maiorino L, Piantedosi V, Cigolari S, Manzione L, *et al.* Vinorelbine (VNR) plus best supportive care (BSC) vs BSC in the treatment of advanced non-small cell lung cancer (NSCLC) elderly patients (pts). Results of a phase III randomized trial. *Ann Oncol* 1998;**9**(Suppl 4):85.

Gridelli C, Maiorino L, Monfardini S. Vinorelbine (VNR) plus best supportive care (BSC) vs BSC in the treatment of advanced non-small cell lung cancer (NSCLC) elderly patients (pts). Results of a phase III randomized trial. *Br J Cancer* 1998;**78**(Suppl 2):3.

Kelly K, Crowley J, Bunn PA, Livingston RB, Gandara DR. Toxicity analysis of SWOG 9509: a randomized phase III trial of paclitaxel plus carboplatin (PC) versus vinorelbine plus cisplatin (VC) in untreated patients with advanced non-small cell lung cancer (NSCLC). A Southwest Oncology Group trial. *Semin Oncol* 1999;**26**(Suppl 18):24.

Kosmidis P, Mylonakis N, Skarlos D, Samantas E, Demopoulos M, Papadimitriou C, *et al.* Paclitaxel (175 mg/m^2) plus carboplatin versus paclitaxel (225 mg/m^2) plus carboplatin in inoperable non-small cell lung cancer: a randomized trial of the Hellenic Cooperative Oncology Group. *Semin Oncol* 1999;**26**(Suppl 2):73.

Kosmidis PA, Mylonakis N, Skarlos D, Samantas E, Demopoulos MA, Papadimitriou C, *et al.* A multicenter randomized trial of paclitaxel (175 mg/m^2) plus carboplatin (6AUC) versus paclitaxel (225 mg/m^2) plus carboplatin (6AUC) in advanced non-small cell lung cancer (NSCLC). *Eur J Cancer* 1999;**35**(Suppl 4):S247.

Lueck HJ, Thomssen C, Kuhn W, Eidtmann H, du Bois A, Moebus V. Randomized multicenter phase III study of epirubicin/paclitaxel versus epirubicin/cyclophosphamide in first-line treatment of metastatic breast cancer: a study of the AGO breast cancer study group. *Semin Oncol* 1999;**26**(Suppl 18):8.

Martinez-Aguillo M, Perez-Calvo J, Aramendia JM, Inoges S, Subira ML, Fernandez-Hidalgo O, *et al.* Prospective randomized study of chemotherapy versus chemotherapy with immunotherapy in stage IIIB or IV non-small-cell lung cancer. *Acta Haematol* 2000; **103**(Suppl 1):55.

Mavroudis D, Papadakis E, Veslemes M, Agelidou M, Toubis M, Agelidou A, *et al.* Paclitaxel–cisplatin–etoposide (TEP) versus cisplatin–etoposide (EP) as first line treatment in small cell lung cancer (SCLC): a preliminary analysis of a multicenter randomized phase III trial. *Eur J Cancer* 1999;**35**(Suppl 4):S247.

May E, Ollivier L, Thiesse P, Seynaeve P, Grau B, Woussen F, *et al.* Experience of an external radiological review committee during a randomized pan-European trial comparing paclitaxel (TAX)/carboplatin (CAR) versus paclitaxel/cisplatin (CIS) in advanced non-small cell lung cancer (NSCLC). *Eur J Cancer* 1999; **35**(Suppl 4):S246.

Pisters KMW, Ginsberg RJ, Putnam JB, Johnson DH, Roberts JR, Kris MG, *et al.* Paclitaxel and carboplatin prior to surgery in early stage non-small cell lung cancer (NSCLC). *Semin Oncol* 1999;**26**(Suppl 18):19.

Sandler A, Denham C, von Pawel J, Cormier Y, Gatzemeier U, Mattson K, *et al.* Significantly higher survival with gemcitabine–cisplatin vs cisplatin in advanced non-small cell lung cancer (NSCLC): final results from an international randomized phase III study. *Ann Oncol* 1998;**9**(Suppl 4):85.

Scagliotti G, Crino L, Conte PF, Rinaldi M, De Marinis F, Gridelli C. Randomized phase III study of gemcitabine/cisplatin versus mitomycin/ifosfamide/cisplatin in advanced non-small cell lung cancer (NSCLC). *Ann Oncol* 1998;**9**:407.

Scagliotti GV, Giaccone G, Postmus PE, Splinter TAW, Van Meerbeeck J, Debruyne C. Teniposide/cisplatin versus paclitaxel/cisplatin in advanced non-small cell lung cancer: interim results of a randomized phase III study of the EORTC Lung Cancer Cooperative Group. *Eur Respir J Suppl* 1996;**9**(Suppl 23):399S–400S.

Skarlos DV, Mylonakis N, Fountzilas G, Samantas E, Dimopoulos MA, Papadimitriou C, *et al.* A multicenter randomized phase III study of paclitaxel (175 mg/m^2) plus carboplatin (AUC 6) versus paclitaxel (225 mg/m^2) plus carboplatin (AUC 6) in advanced non-small cell lung cancer (NSCLC). *Semin Oncol* 1999; **26**(Suppl 18):25.

Souquet PJ, Huat TE, Pereira JR, Van Klaveren R, Price A, Gatzemeier U. GLOB 1: a phase III trial comparing the combination of vinorelbine and cisplatin (NP) versus vinorelbine, ifosfamide and cisplatin (NIP) in metastatic non small cell lung cancer patients (NSCLC). *Ann Oncol* 2000;**11**(Suppl 4):4.

Vansteenkiste J, Vandebroek J, Nackaerts K, Deschepper K, Van Kerckhoven W, Vandeurzen K, *et al.* Gemcitabine monotherapy (GEM) versus cisplatin-based chemotherapy in symptomatic advanced NSCLC: a prospective randomised comparison of symptom control. *Lung Cancer* 2000;**29**:48.

Vokes EE, Leopold KA, Herndon JE, Crawford J, Perry MC, Miller AA, *et al.* CALGB Study 9431: a randomised phase II study of cisplatin with gemcitabine or paclitaxel or vinorelbine as induction chemotherapy (Ind CT) and concomitant chemoradiotherapy (XRT) for unresectable stage III non-small cell lung cancer (NSCLC). *Lung Cancer* 2000;**29**:49.

Appendix 1

Rapid review methods from the research protocol

The methods below were approved by NICE at the start of the review.

Search strategy

The following electronic databases were the first ones searched: the Cochrane Library, MEDLINE, EMBASE, DARE, NHS EED, Physicians Data Query (PDQ), Science Citation Index, Medical Research Council (MRC) Trials database and National Research Register. Some manufacturers provided lists of trials that they intended to cite, and this was a useful check on the comprehensiveness of our searches. Searches were limited to English language. Bibliographies were checked for relevant studies.

Experts, including the Cochrane Lung Cancer Group, were contacted for advice and peer review, and to identify additional published and unpublished references.

Searches for QoL data used the restricted term 'health-related QoL'.

Inclusion and exclusion criteria

Interventions include the four drugs, used alone or in combination with other drugs, for NSCLC and SCLC. Other tumours, such as mesothelioma of the lung, were not included, nor were lung metastases from primaries outside the lung. The use of drugs as adjuvant therapy before surgery, or in combination with radiotherapy was not included, following review in liaison with NICE of data from manufacturers.

Initial searches were for systematic reviews, RCTs, QoL studies and economic appraisals. Office of National Statistics data were used for incidence and mortality rates.

Data extraction from trials was done by one person and checked by at least one other.

Quality assessment for RCTs was performed using the Jadad system. The CRD assessment criteria were to be used for systematic reviews. A list of appraisal questions for economic studies is appended.

Methods of analysis

It was agreed that meta-analysis would be considered if the volume of evidence would support this, but there were insufficient studies of the same drug against the same comparator.

Possible subgroups and comparators were expected to be:

- first-line treatment for good prognosis SCLC – CAV, PE or CAV/PE
- first-line treatment for poor prognosis SCLC – PE, or VP-16 and platinum
- chemotherapy for NSCLC after radical radiotherapy – MIC or MVP
- palliative chemotherapy for NSCLC versus BSC
- second-line chemotherapy for NSCLC versus BSC.

However, only NSCLC was found to be relevant for these drugs at present.

Methods for estimating QoL, costs and cost-effectiveness

QoL information for estimation of QALYs was obtained from the literature. Experts were to be consulted regarding unpublished work. Evidence on patients' views was sought from the literature and from the Roy Castle Foundation.

Costs were based partly on published NHS data, partly from scientific publications and partly from unpublished SHPIC work that investigated the costs of care of a sample of patients with lung cancer. Data from overseas publications were used with caution. Only NHS acute sector costs were estimated. Consideration was given to average and marginal costs of hospital care.

When appropriate, cost per QALY was to be estimated by combining effectiveness data from the trials, QoL information and cost estimates. Sensitivity analyses were performed to determine how robust the estimates are relative to the assumptions made.

Appendix 2

Sources of information, including databases searched and search terms

The following databases were searched for published studies, recently completed studies and ongoing research.

Database	Most recent date searched
The Cochrane Library (Issue 3, 2000)	14 Jul 2000
MEDLINE (SilverPlatter®)	18 Jul 2000
'PreMEDLINE' (Internet – Grateful Med)	19 Jul 2000
EMBASE (SilverPlatter)	21 Jul 2000
EconLit	19 Jul 2000
CRD databases (NHS EED, DARE, HTA)	25 Jul 2000
CancerLit	25 Jul 2000
National Research Register (Internet resource)	25 Jul 2000
MRC Clinical Trials (Internet resource)	2 Aug 2000
Cancer Trials (Internet resource)	4 Aug 2000
Current Controlled Trials (Internet resource)	8 Aug 2000

The following search terms were used to identify relevant articles.

For searches on NSCLC and SCLC	
Medical Subject Headings (MeSH)	**EMTREE Headings**
"Lung-Neoplasms"	"Lung-Carcinoma"
"Carcinoma-Small-Cell"	"Neuroendocrine-Tumor"
"Neuroendocrine-Tumors"	"Lung-Squamous-Cell-Carcinoma"
"Carcinoma-Non-Small-Cell-Lung"	"Adenocarcinoma"
"Carcinoma-Squamous-Cell"	"Giant-Cell"
"Adenocarcinoma"	"Large-Cell-Carcinoma"
"Carcinoma-Adenosquamous"	
"Carcinoma-Giant-Cell"	
"Carcinoma-Large-Cell"	

Free-text words used for searches on NSCLC and SCLC		
lung* in ti, ab	sclc in ti, ab	adeno-carcin* in ti, ab
carcin* in ti, ab	small cell in ti, ab	adenocarcin* in ti, ab
tumor* in ti, ab	neuroendocrin* in ti, ab	adeno-squamous* in ti, ab
tumour* in ti, ab	nsclc in ti, ab	adenosquamous* in ti, ab
neoplasm* in ti, ab	non small cell in ti, ab	giant cell in ti, ab
cancer* in ti, ab	squamous in ti, ab	large cell in ti, ab
oat cell in ti, ab	epidermoid* in ti, ab	
bronch* in ti, ab	spindle cell in ti, ab	

Free-text and EMTREE headings used for searches on chemotherapy drugs	
Gemcitabine	***Paclitaxel***
Gemcitabin*	Paclitax*
Gemcytabin*	Paclitac*
Gemzar	Paxene
difluorodeoxytidine	Anzatax
Vinorelbine	***Docetaxel***
Vinorelbin*	Docetaxel*
Vinoralbin*	Docetaxol
5′ noranhydrovinblastine	Navelbine
Taxotere	
MeSH/EMTREE headings and free-text words used for searches on costs	
"Economics"	cost*
"Costs-and-Cost-Analysis"	price*
"Cost-Benefit-Analysis"	econom*
"Economics-Nursing"	Pharmoeconom*
"Economics-Medical"	quality-of-life
"Economics-Medical"	QOL
"Economics-Pharmaceutical"	HRQOL
"Economics-Hospital"	QALY*
"Quality-of-Life"	LYG
"Quality-of-Adjusted-Life-Years"	
All searches (with the exception of those attempting to identify costs of therapy) were restricted to reviews, meta-analyses, and randomised or controlled clinical trials. The number of articles identified and subsequently reviewed for this report are shown in Figures 2 *and* 3	

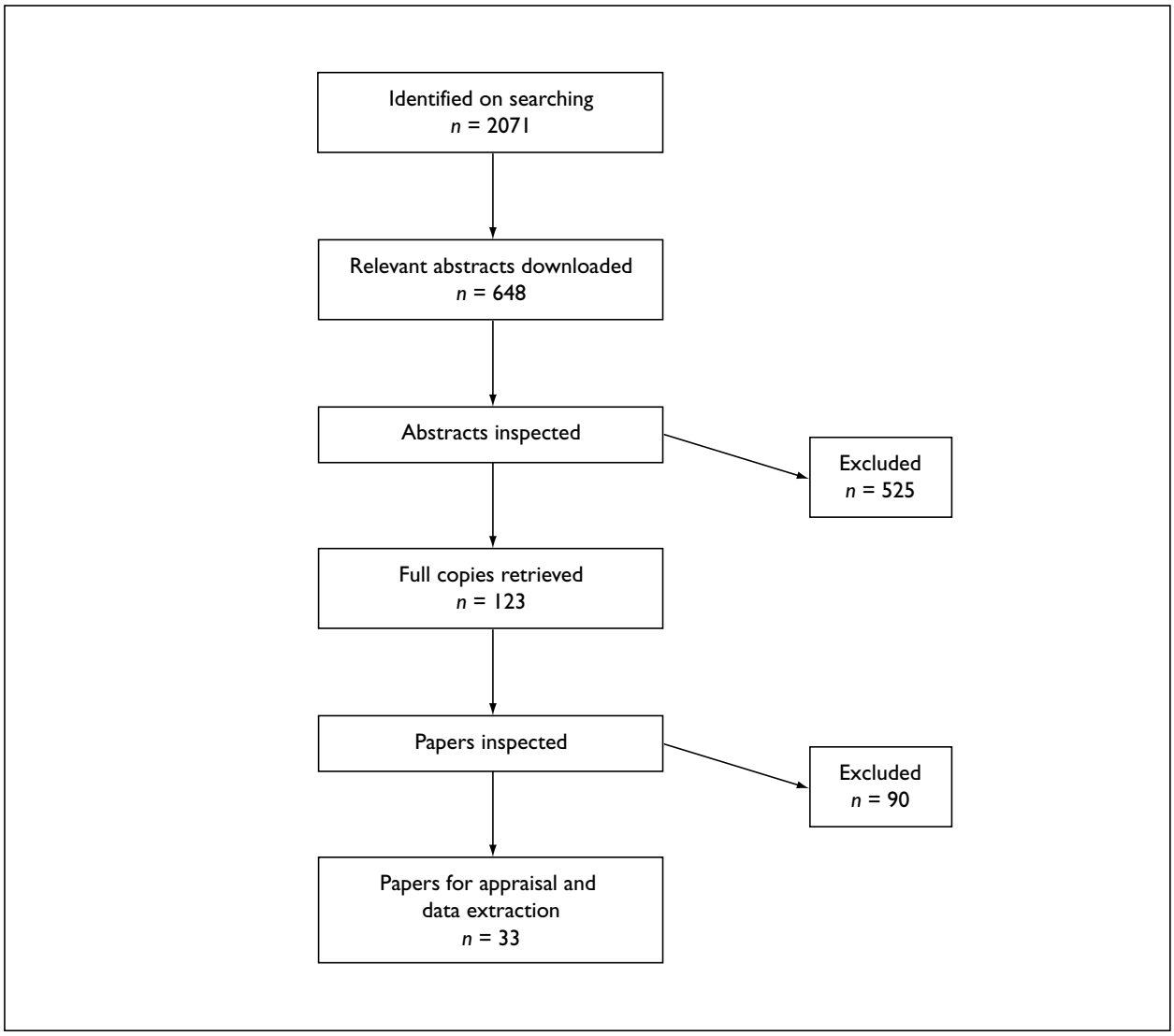


FIGURE 2 *Flow chart of identification and inclusion of studies for reviews, meta-analyses, and randomised or controlled trials. (Note: An update search of all databases used during the original literature search revealed one further RCT [Gatzemeier and co-workers, 2000[34]] that was included and critically appraised)*

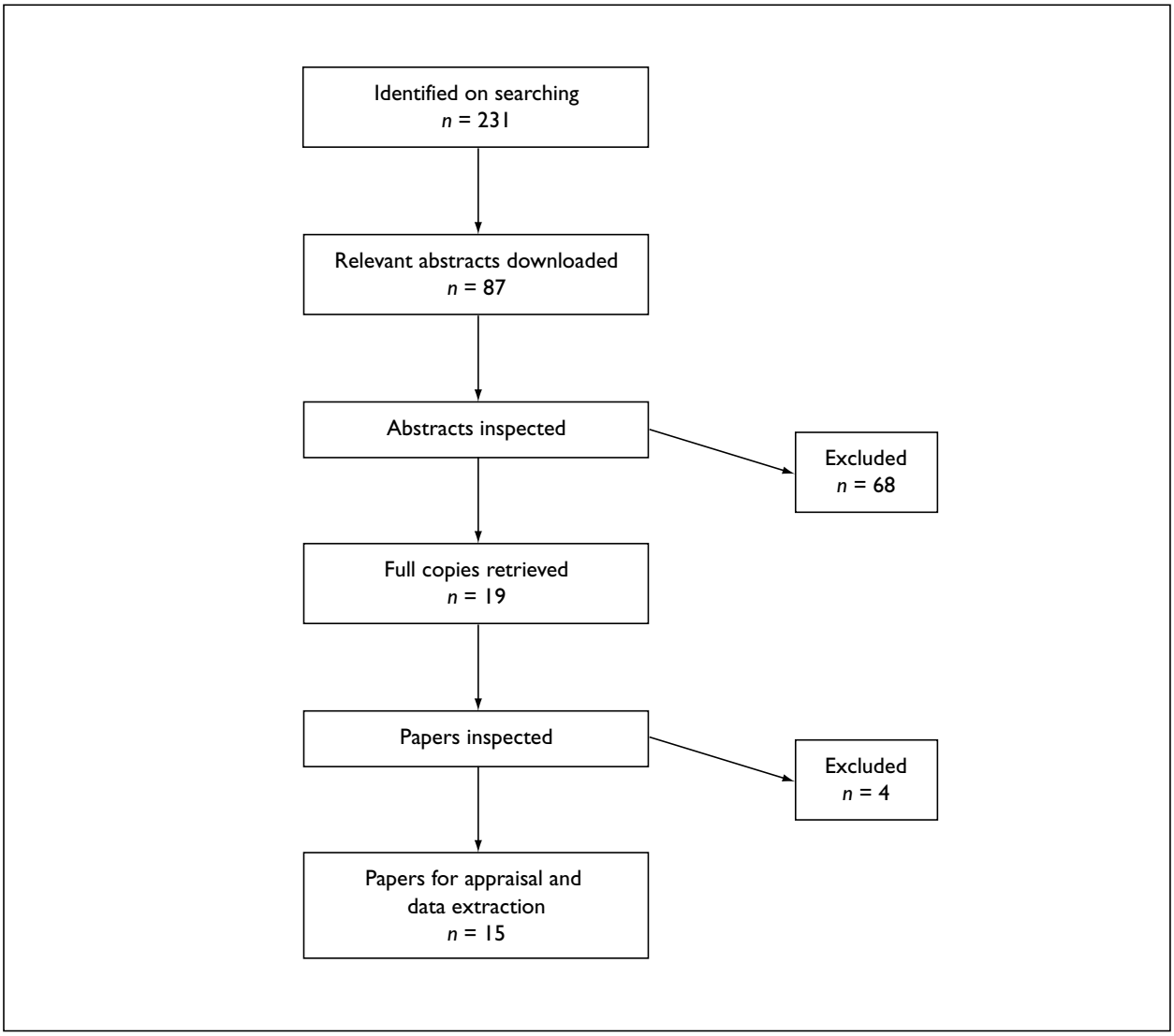


FIGURE 3 *Flow chart of identification and inclusion of studies for economic evaluations*

Appendix 3

Instrument used to measure the likelihood of bias in RCTs (Jadad quality scale)[19]

Questions to assess the likelihood of bias

1. Was the study described as randomised (this includes the use of the words such as randomly, random and randomisation)?
2. Was the study described as double-blind?
3. Was there a description of withdrawals and dropouts?

Scoring the items

Give a score of either 1 point for each 'yes' or 0 points for each 'no'. There are no in-between marks.

Give 1 additional point if:

- for question 1, the method used to generate the sequence of randomisation was described, and it was appropriate (e.g. table of random numbers, computer generated) and/or
- for question 2, the method of double-blinding was described, and it was appropriate (e.g. identical placebo, active placebo, dummy).

Deduct 1 point if:

- for question 1, the method used to generate the sequence of randomisation was described, but it was inappropriate (e.g. patients were allocated alternately or according to, for example, date of birth or hospital number) and/or
- for question 2, the study was described as double-blind, but the method of blinding was inappropriate (e.g. comparison of tablet vs injection with no double dummy).

Guidelines for assessment

1. Randomisation

A method to generate the sequence of randomisation will be regarded as appropriate if it allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which treatment was next. Methods of allocation using date of birth, date of admission, hospital numbers or alternation should not be regarded as appropriate.

2. Double-blinding

A study must be regarded as double-blind if the word 'double-blind' is used. The method will be regarded as appropriate if it is stated that neither the person doing the assessments nor the study participant could identify the intervention being assessed, or if in the absence of such a statement the use of active placebos, identical placebos or dummies is mentioned.

3. Withdrawals and dropouts

Participants who were included in the study but did not complete the observation period or were not included in the analysis must be described. The number and the reasons for withdrawal in each group must be stated. If there were no withdrawals, it should be stated in the article. If there is no statement on withdrawals, this item must be given no points.

Appendix 4

Quality assessment scales for systematic reviews[20]

Criteria for assessing the quality of systematic reviews

Systematic reviews will be examined to determine how many of the following criteria for methodological quality they met.

1. Does the review answer a well-defined question?
A good review should focus on a well-defined question, making the objectives of the review easy to understand. The most important components in a review question include the target population, healthcare intervention and outcomes of interest.

2. Was a substantial effort made to search for all the relevant literature?

3. Are the inclusion/exclusion criteria reported and are they appropriate?
Criteria for the inclusion of individual studies in a review have two major dimensions: relevance and validity. A relevant study should be useful to answer review questions in terms of patients, intervention and outcomes. The validity issue is related to the methodological standard of an individual study.

4. Is the validity of included studies adequately assessed?

5. Is sufficient detail of the individual studies presented?
Details of the individual studies included in a review include study design, sample size in each study group, patient characteristics, description of interventions, settings, outcome measures, follow-up, dropout rate, effectiveness results and side-effects. The importance of the study details may differ for different review topics.

6. Have the primary studies been combined or summarised appropriately?
If at least four of the criteria are met, the paper will be considered to be of good quality.

Appendix 5

Summary of evidence of effectiveness of docetaxel in lung cancer

Reference and design	Intervention	Participants	Outcome measures
Shepherd *et al.*, 2000[21] (International) Phase III, multicentre, randomised trial Jadad score: 2/3	***DOC arm*** DOC 100 mg/m^2 (iv) every 21 days for first half of trial; DOC 75 mg/m^2 (iv) every 3 weeks for second half of the study ***BSC arm*** BSC; therapy determined by treating physician (e.g. antibiotics, analgesic drugs, transfusions, palliative therapy) ***Length of treatment*** Treatment continued in DOC arm until disease progression or unacceptable toxicity developed (evaluation every 3 weeks) ***Other interventions used*** Oral dexamethasone 8 mg b.d. for 5 days, 24 hours before DOC (10 doses)	*n* = 204 DOC 100 mg/m^2 arm: *n* = 49 DOC 75 mg/m^2 arm: *n* = 55 BSC arm: *n* = 100 67% of patients were male 80% of patients had stage IV NSCLC 76% of patients had PS of 0 or 1 ***Characteristics of target population*** All patients must have received prior treatment with platinum-containing (CDDP/CBDCA) chemotherapy regimen (1+ regimens not including taxanes) Histological/cytological diagnosis of NSCLC (stage IIIA/B or IV) Median age: DOC arm, 61 years (range, 37–76 years); BSC arm, 61 years (range, 28–77 years) PS of 2 or lower (ECOG scale) See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival time: Calculated from the date of randomisation to the date of death. Survival time was censored for loss of contact or initiation of antitumour therapy (including subsequent chemotherapy, immunotherapy or surgery) ***Secondary end-points*** Objective tumour response: WHO response criteria, responses confirmed in 4+ weeks after initial documentation Duration of response: Calculated from the date of randomisation to the date of documentation of disease progression QoL: LCSS used for North American patients, EORTC QLQ-C30 used for European patients ***Length of follow-up*** Patients in both arms were assessed every 3 weeks. Complete medical history and physical examination, with documentation of weight and ECOG PS. Vital signs and toxicities also evaluated. QoL questionnaires completed every 3 weeks. Chest X-ray every 3 weeks, and scans repeated every 6 weeks to document response and disease progression. In the BSC arm, patients were evaluated every 3 weeks for the first 18 weeks and then every 6 weeks See *General comments* below

Results

- *Survival*: Survival curves shown, DOC curve above BSC curve (for both overall DOC arm and DOC 75 mg/m^2 group, but comparable for DOC 100 mg/m^2 group). Median survival of DOC arm was 7 months (range, 5.5–9.0 months) and of BSC arm 4.6 months (range, 3.7–6.0 months) (log-rank test, $p = 0.047$). One-year survival rates were 29% for DOC and 19% for BSC. Median survival of 75-mg/m^2 DOC arm was 7.5 months compared with 4.6 months in BSC arm (log-rank test, $p = 0.01$). One-year survival was 37% for DOC arm, 12% for BSC arm (χ^2 test, $p = 0.003$). Median survival of 100 mg/m^2 DOC arm was 5.9 months, with 1-year survival 19%. No difference was evident between 100-mg/m^2 DOC arm and BSC arm ($p = 0.78$). Cox modelling showed that, after adjustment for prognostic factors (PS, tumour stage, number of organs involved, number of prior chemotherapy regimens and total baseline score on LCSS), treatment with DOC 75 mg/m^2 had a significant impact on survival (hazards ratio, 0.484; $p = 0.004$)
- *Response*: No patients achieved complete response. Six (5.8%) patients in the overall DOC arm achieved partial response (3 patients in DOC 100 mg/m^2 and 3 patients in DOC 75 mg/m^2). Five responders had a PS of 1 (1 responder with PS of 0). Time to progression for DOC arm patients was significantly longer overall (10.6 weeks vs 6.7 weeks, $p = 0.001$). This effect was also seen at both dose levels (DOC 100 mg/m^2, $p = 0.037$; DOC 75 mg/m^2, $p = 0.004$)
- *Duration of response:* The median duration of response was 26.1 weeks (range, 23.7 to 31.0+ weeks)
- *QoL:* Results of questionnaire responses were reported in another article. However, all QoL parameters favoured DOC-treated patients, and differences were statistically significant for pain and fatigue scales ($p = 0.006$ and 0.06, respectively). PS mean change in DOC patients was 0.56, BSC 0.8 ($p = 0.11$). Seven per cent of DOC patients experienced more than 10% weight loss in comparison with 15% of BSC patients ($p = 0.07$). Fewer tumour-related medications were used for DOC than BSC ($p = 0.02$). Use of morphine/morphine-equivalent medications for pain was 32% for DOC vs 49% for BSC ($p = 0.01$). Non-morphine analgesic use was 39% for DOC vs 55% for BSC ($p = 0.03$). Medications for tumour-related indications other than pain: 30% for DOC vs 49% for BSC ($p < 0.01$). Radiotherapy (at least one treatment during study treatment or follow-up) was required by fewer DOC patients (26% for DOC vs 37% for BSC; $p = 0.09$)

continued

***Results* contd**

- *Haematological toxicity*: DOC 100 mg/m^2 is slightly more toxic than 75 mg/m^2. Severe grade 3 neutropenia occurred in 86% and 67% of patients being treated with DOC, 100 mg/m^2 and 75 mg/m^2, respectively. Grade 3/4 anaemia occurred in 16.3% and 5.5% of patients receiving DOC, 100 mg/m^2 and 75 mg/m^2, respectively. Grade 3/4 febrile neutropenia occurred in 22.4% and 1.8% of DOC patients. Thrombocytopenia occurred in 1% of all DOC patients. Overall, septic deaths occurred in 2.9% of patients; however, none of these were from the 75-mg/m^2 arm, whereas 6.1% of the 100-mg/m^2 arm suffered septic death
- *Non-haematological toxicity*: All groups observed toxicities. Nausea and vomiting were equally distributed across all groups, but diarrhoea occurred in only the DOC arm. Infection occurred in 34% of chemotherapy patients, but 21% of BSC patients suffered this too

Comments

Methodological comments

- Groups were randomised, but no indication as to how they were randomised. No placebo control
- Control arm was BSC. Patients were stratified on the basis of their ECOG PS and their best response to platinum-based therapy (progression vs no change, partial response or complete response
- Baseline characteristics, in terms of age, gender, stage of NSCLC and PS were equally distributed
- No specification of blinding. No specification of equal treatment
- Using ITT analysis, there was an overall response rate of 5.8%. Cox modelling was used to determine the impact of several prognostic variables on survival
- Concern exists for the Cox modelling used to determine the impact of prognostic variables on survival. The authors report that one of the significant variables adjusted in the modelling (LCSS) was only applicable to the North American population. They did not report if the modelling was based on a subset of the total DOC population
- Sample size of 104 randomised to DOC arm and 100 to BSC. Sample size of 100 patients per group estimated on projected median survival of 7 months in the DOC group and 4 months in the BSC group, on basis of log-rank test with an alpha level of 5% (2-sided) and a power of 90% to compare groups
- Concern exists that comparisons between DOC 100 mg/m^2 or DOC 75 mg/m^2 and BSC are not sufficiently powered to show a reliable treatment effect
- Uncertainty exists surrounding how the results of the responses to the QoL questionnaires (LCSS and EORTC QLQ-C30) were combined. Concern exists that significant differences were reported between QoL variables in the article without further documentation
- One patient was excluded from ITT analysis because of unconfirmed NSCLC (results reported in this appraisal based on ITT)
- Eligibility and exclusion criteria were clearly defined. Primary outcome was survival time, with secondary outcomes toxicity, QoL and response assessment
- Survival time was measured from the date of randomisation to date of death. Toxicity was measured using WHO code, QoL measured via the LCSS in North America and the EORTC QLQ for lung cancer was used in Europe. PS was assessed using ECOG scales. Objective tumour response and response duration in DOC arm were assessed using WHO response criteria. Response duration was calculated from date of randomisation to date of disease progression

General comments

- Date (length of study time) to "interim safety-data monitoring" when dose of DOC was reduced (amendment to original protocol) from 100 mg/m^2 to 75 mg/m^2 for randomised patients was not stated
- Eligibility and inclusion criteria were clearly defined. Additional inclusion criteria included: patients with both measurable and evaluable lesions permitted; adequate haematological parameters; serum creatinine level of 2.0 mg d/l; no symptomatic or uncontrolled brain metastases or peripheral neuropathy greater than NCI grade 2; patients still eligible if received prior radiation therapy, provided that ≤ 25% of their total bone marrow had been irradiated, but had to wait 30 days before entry to study; required to wait 21 days before entry to study after treated with any chemotherapy, immunotherapy or biologic systemic anticancer therapy
- All patients underwent physical examination and complete medical history (within 7 days of study entry), including full neurological examination and documentation of ECOG PS). Also, complete haematological and biochemical testing and an ECG and baseline chest X-ray were performed. Within 3 weeks before therapy, QoL questionnaire was completed, and clinically indicated scans including CT scans of the brain, thorax and upper abdomen (or abdominal ultrasound) and radionuclide bone scans were performed
- *Conflicts of interest*: None stated

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	Not possible due to comparison with BSC
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	One member of the DOC group was excluded from ITT analysis due to unconfirmed NSCLC 1/104 patients (1%) excluded from overall DOC arm 1/49 patients (2%) excluded from 100-mg/m^2 DOC arm

Reference and design	Intervention	Participants	Outcome measures
Roszkowski *et al.*, 2000[22] Phase III, open-label, multicentre, randomised trial (International) No funding information provided Jadad score: 2/3	***DOC arm*** DOC 100 mg/m^2 (iv) every 3 weeks ***BSC arm*** BSC as judged by treating physician. Included use of antibiotics, analgesics, transfusions or any other symptomatic treatment medically indicated. No chemotherapy or other systemic anticancer therapy, except radiotherapy See *General comments* below ***Length of treatment*** DOC continued until signs of progressive disease, unacceptable side-effects or when estimated maximum benefit obtained ***Other interventions used*** DOC patients given 8 mg dexamethasone pre-treatment, G-CSF, antiallergic and anti-emetics (except steroids) permitted except first cycle	*n* = 207 DOC arm: *n* = 137 BSC arm: *n* = 70 82% of patients were male 80% of patients had PS of 0 or 1 47% of patients had stage IV NSCLC 52% of patients had squamous cell carcinoma 29% of patients had adenocarcinoma ***Characteristics of target population*** Histologically/cytologically unresectable stage IIIB and IV NSCLC Patients aged between 18 and 75 years Median age: DOC arm, 59 years (range, 36–75 years); BSC arm, 60 years (range, 31–74 years) Uni- or bi-dimensionally measurable disease of which at least one area had not been subject to prior irradiation PS of 0 to 2 (WHO PS) Adequate organ function No prior chemotherapy or immunotherapy (even as neo-adjunctive or adjunctive therapy) Previous radiotherapy limited to ≤ 25% of bone marrow and completed > 4 weeks prior to enrolment, patients fully recovered from toxic effects See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculated from the date of randomisation until the date of death (median, 1- and 2-year survival rates) ***Secondary end-points*** Time to progression: Calculated from the date of randomisation to the date of death Response: WHO criteria evaluated every 3 cycles until progression QoL: EORTC QLQ-C30 completed at baseline, then every second cycle for DOC arm and every 6 weeks for BSC arm, then every 3 months of follow-up Toxicity: Graded according to NCI criteria. Adverse events not gradable by NCI criteria; used COSTART. Evaluated at the end of each cycle ***Length of follow-up*** All patients were followed for 36 months

Results

- *Survival*: Survival curves shown, DOC arm above BSC arm (showed a separation after 6 months of follow-up, sustained over time and were significantly different; $p = 0.026$ for the log-rank test and $p = 0.04$ for log-rank test stratified on extent of disease). Median survival with DOC was 6.0 months (95% CI, 5.0 to 8.0 months) vs 5.7 months (95% CI, 4.4 to 6.8 months) with BSC. One-year survival: 25% in DOC arm vs 16% in BSC. Two-year survival: 12% in DOC arm vs 0% in BSC arm (none survived beyond 20 months). In total, 27 patients (16 in DOC arm, 11 in BSC arm) received off-protocol chemotherapy – removing these led to a similar conclusion for the overall survival curve
- *Time to progression*: Significantly longer in DOC arm (12.6 weeks vs 8.9 weeks with BSC)
- *Response*: Response rate in the ITT population was 13.1% (95% CI, 7.5% to 18.8%) and 19.6% (95% CI, 12.0% to 29.1%) in the evaluable-for-response population. There were 18 confirmed responses, including 2 patients with complete response. Median duration of response was 37.1 weeks (95% CI, 30.9 to 69.9 weeks). In the ITT population, 31.4% of the DOC arm achieved disease stabilisation; in the evaluable-for-response group, the rate was 42.4%
- *QoL*: Performed on ITT population. Emotional functioning was significantly in favour of the DOC arm in three of the four methods of analysis – longitudinal mixed model ($p = 0.01$), pattern mixed model ($p = 0.04$) and worst score ($p = 0.01$) – but not significant in the AUC analysis. No statistical significance in difference was evident between global health status/QoL and physical functioning scores for the two arms. Using AUC analysis, there was a significant improvement in nausea/vomiting ($p = 0.04$), pain ($p < 0.0001$) and dyspnoea ($p = 0.02$) in the DOC arm. Using sensitivity analysis, statistically significant differences were found (using mixed model, pattern mixture model and worst score) in favour of the DOC arm for all measures of pain and dyspnoea except nausea/vomiting. None of the other dimensions were significant, but all favoured the DOC arm except diarrhoea
- *Haematological toxicity*: Patients in the BSC arm had a higher incidence of pulmonary events, neurocortical events and pain. Patients in the DOC arm had a higher incidence of asthenia, infection and neurosensory events. Grades 3 and 4: neutropenia (28% vs 0%), leucopenia (22% vs 0%), anaemia (5% vs 3%), asthenia (28% vs 23%) and infections (11% vs 3%) for DOC and BSC arms, respectively. In the DOC arm, 5 (3.6%) patients had febrile neutropenia. Two patients in the DOC arm died from toxicity and 1 other from a combination of malignant disease and infection. No information on statistical significance was detailed for any of the above

continued

***Results* contd**

- *Non-haematological toxicity*: Grades 3 and 4: pulmonary (23% vs 35%), pain (21% vs 33%) and nausea (5% vs 1%) for DOC and BSC arms, respectively. No information on statistical significance was detailed for any of the above. Patients in the DOC arm required less use of opiate analgesics and non-opiate analgesics (both $p < 0.0001$), tumour-related medications other than for pain and palliative radiotherapy (both at $p < 0.01$). DOC arm patients also had a higher use of anti-infective drugs ($p = 0.027$)
- *Serious adverse events*: Hospitalisation for adverse events was required for 51% (70) of patients in the DOC arm and 30% (21) in the BSC arm. Of the DOC patients, 30% had serious events as a consequence of treatment; 19% (26) of patients in the DOC arm dropped out of the study because of adverse events

Comments

Methodological comments

- Trial stated as randomised centrally, but there was no explanation of why patients were randomised as 2:1 in favour of DOC arm (after staging). One patient randomised to BSC arm received DOC by accident. No indication of proportion of patients recruited was provided. No placebo control
- No significant differences in baseline characteristics of patients between the two groups
- Blinding not performed, open-label study
- ITT group was defined as all randomised patients analysed in the group to which they were allocated; should be all randomised patients. Patients receiving 2+ cycles with at least one follow-up visit were considered evaluable for response, provided that the tumour lesions were properly assessed per protocol. CIs given for survival and response. Safety data analysed in all patients in the treatment group they actually received (i.e. not ITT analysis)
- Sample size of 200 patients was necessary to show survival advantage (in DOC arm) with 85% power and alpha level of 5. Survival and time to progression were calculated in the two treatment groups by the Kaplan–Meier method. Two groups were compared using log-rank test. Response rates were provided with the exact 95% CI, in the ITT and evaluable-for-response populations. Global health status/QoL, physical functioning and emotional scores were considered as primary QoL dimensions and thus analysed using four methods: a longitudinal mixed model, worse score, AUC analysis, and pattern mixture model based on completers and non-completers. Completers were defined as patients with follow-up data for QoL after treatment period 5. Non-completers were defined as patients with follow-up data for QoL at or before treatment period 5. Worse score during study and AUC were compared using a non-parametric Wilcoxon test
- In the DOC arm, 26 (19%) patients dropped out as a result of adverse events. No data on any withdrawals in the BSC arm

General comments

- Deviations from protocol: 27 patients received off-protocol chemotherapy (16 in DOC arm and 11 in BSC arm)
- Large proportion of missing QoL questionnaires in BSC arm at baseline (41% vs 7% in DOC arm) attributed to inability of investigators to differentiate between baseline and period 1
- Clearly defined inclusion and exclusion criteria. Additional inclusion criteria included: no presence or history of symptomatic CNS metastases; no pre-existing neuromotor or neurosensory toxicity (NCI grade 2); no serious illness or medical conditions (e.g. infection requiring antibiotics, active cardiac disease); no past or present history of neoplasm, except curatively treated non-melanoma skin cancer, cervical carcinoma *in situ* or other cancer curatively treated, with no evidence of the disease for the past 5 years
- Appropriate outcome measures were used
- Prior to treatment, medical history and physical examination (including clinical tumour assessment, WHO PS, vital signs, concomitant medications, baseline signs and symptoms), neurological examination, and haematological and serum chemistry evaluations were performed. ECG and radiological examinations were performed to document lesions. For DOC patients, physical examination (including clinical tumour assessment, WHO PS, vital signs, concomitant medications and adverse events) and haematological evaluations were performed after every cycle; chemistry evaluations were performed every other cycle, and X-rays and CT scans every 3 cycles until progression. For BSC patients, physical examination were performed (as for DOC patients) every 3 weeks, haematological and chemistry evaluations every 6 weeks, and X-rays and CT scans every 9 weeks. Both arms had neurological and bone scans, as clinically indicated
- *Conflicts of interest*: No drug industry involvement was cited, but three authors were employed by drug company

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	Not possible due to comparison with BSC
Was there a description of withdrawals and dropouts?	1 concern over definition of ITT group
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	26 (19%) of patients in the DOC arm withdrew due to adverse events; no data on BSC arm

Reference and design	Intervention	Participants	Outcome measures
Fossella *et al*., 2000[23] (USA) Phase III, open-label, multicentre, randomised trial Funding: Supported by Rhône-Poulenc Rorer Jadad score: 2/5	***DOC (100) arm*** DOC 100 mg/m^2 (iv) every 3 weeks ***DOC (75) arm*** DOC 75 mg/m^2 (iv) every 3 weeks ***VNB or IFOS arm*** Participating investigators free to choose either VNB or IFOS. VNB 30 mg/m^2 (iv) on days 1, 8 and 15 of each 3-week cycle or IFOS 2 mg/m^2 (with standard dose mesna) on days 1, 2 and 3 of each 3-week cycle ***Length of treatment*** Minimum of 2 cycles. Patients with response or stable disease continued treatment for at least 6 cycles, unless disease progressed. Patients could receive more than 6 cycles if they were responding and stable, and they were seen to been achieving clinical benefit, as determined by the treating physician ***Other interventions used*** Oral dexamethasone (8 mg) every 12 hours, for 5 doses, starting 24 hours before DOC infusion. Parenteral antibiotics, prophylactic filgrastim, mesna for IFOS arm, prophylactic antiemetics	*n* = 373 DOC (100) arm: *n* = 125 DOC (75) arm: *n* = 125 VNB/IFOS arm: *n* = 123 (VNB, *n* = 89; IFOS, *n* = 34) 65% of patients were male 17% of patients had PS of 2 50% of patients had adenocarcinoma histology 89% of patients had stage IV NSCLC ***Characteristics of target population*** Median age: DOC (100) arm, 60 years; DOC (75) arm, 60 years; VNB/IFOS arm, 60 years Locally advanced or metastatic NSCLC that had progressed during or after one or more platinum-based regimens Minimum of 21 days must have elapsed since prior chemotherapy PS of ≤ 2 (ECOG scale) No restrictions on the number of prior chemotherapy regimens, the amount of prior chemotherapy or the agents used (which may have included PAX) See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculated from the date of randomisation until the date of death ***Secondary end-points*** Response rate: WHO criteria, evaluated before each therapy cycle Response duration: Calculated from the time of randomisation to the first objective evidence of tumour response Time to progression: Calculated from the time of randomisation to the first objective evidence of tumour response Toxicity: NCI common toxicity criteria, evaluated before each therapy cycle QoL: LCSS questionnaire, completed at baseline, before and at the end of each therapy cycle, and at follow-up every 2 months See *General comments* below ***Length of follow-up*** On removal from the study, patients were observed every 2 months until death, to assess adverse events, QoL, disease status and survival

Results

- *Survival*: Survival curves shown, DOC (100) and DOC (75) curves comparable with VNB/IFOS curve until approximately week 18 (VNB/IFOS curve under other treatment groups). Median survival: DOC (100), 5.5 months; DOC (75), 5.7 months; VNB/IFOS, 5.6 months. One-year survival rates: DOC (100), 21% (95% CI, 14% to 28%); DOC (75), 32% (95% CI, 23% to 40%); VNB/IFOS, 19% (95% CI, 12% to 26%). The DOC (75) arm's 1-year survival is significantly greater than that of the VNB/IFOS arm (chi-squared test, $p = 0.025$). Differences in 1-year survival rates between the DOC (100) and VNB/IFOS arms were not significant
- *Survival after censoring for patients who received subsequent chemotherapy*: DOC (100), 6.6 months; DOC (75), 5.8 months; VNB/IFOS, 5.4 months. One-year survival: DOC (100), 32% (95% CI, 22% to 43%); DOC (75), 32% (95% CI, 20% to 44%); VNB/IFOS, 10% (95% CI, 1% to 18%). The 1-year survival rates were significantly greater for both DOC arms compared with VNB/IFOS arm. DOC (100) vs VNB/IFOS ($p = 0.001$), DOC (75) vs VNB/IFOS ($p = 0.002$) and DOC (100 + 75) vs VNB/IFOS ($p < 0.0001$)
- *Response*: Objective response assessed in 358 patients with confirmed NSCLC, who received at least one chemotherapy infusion after randomisation. Partial response: DOC (100), 10.8% (95% CI, 6.1% to 18.1%); DOC (75), 6.7% (95% CI, 3.1% 13.1%); VNB/IFOS, 0.8% (95% CI, 0.0% to 5.3%). Both the response rates for DOC (100) and DOC (75) were significantly higher than the response rate for VNB/IFOS (Fisher's exact test, $p = 0.001$ and $p = 0.036$, respectively, and $p = 0.002$ for both DOC arms combined)
- *Median response duration*: DOC (100), 7.5 months; DOC (75), 9.1 months
- *Median time to progression*: DOC (100), 8.4 weeks (95% CI, 6.7 to 11.0 weeks); DOC (75), 8.5 weeks (95% CI, 6.7 to 11.0 weeks); VNB/IFOS, 7.9 weeks (95% CI, 6.9 to 11.0 weeks). Favoured treatment with DOC: log-rank test, $p = 0.44$ for DOC (100) vs VNB/IFOS, and $p = 0.46$ for both DOC groups combined vs VNB/IFOS. Time to progression curves shown
- *Progression-free survival at 26 weeks (6 months)*: DOC (100), 19%; DOC (75), 17%; VNB/IFOS, 8%. This difference was significant: chi-squared test, $p = 0.013$ for DOC (100) and $p = 0.031$ for DOC (75) compared with VNB/IFOS arm, and $p = 0.005$ when both DOC doses were combined

continued

***Results* contd**

- *Haematological toxicity*: Grade 4 toxicities: neutropenia (77% vs 54% vs 31%; DOC(100) and DOC(75) significantly different from VNB/IFOS), febrile neutropenia (12% vs 8% vs 1%; DOC(100) and DOC(75) significantly different from VNB/IFOS), infection (3% vs 0% vs 1%), anaemia (1% vs 0% vs 2%), thrombocytopenia (1% vs 2% vs 0%), use of filgrastim (% of cycles, 28% vs 7% vs 3%; DOC(100) significantly different from DOC (75) and VNB/IFOS); for DOC (100), DOC (75) and VNB/IFOS, respectively
- *Non-haematological toxicity*: Grade 3 or 4: neurosensory (6% vs 1% vs 3%), diarrhoea (3% vs 2% vs 2%), asthenia (17% vs 12% vs 11%), fluid retention (4% vs 1% vs 2%), nausea (7% vs 3% vs 6%), vomiting (7% vs 1% vs 4%), discontinuation due to toxicity (13% vs 7% vs 4%), treatment-related death (2% vs 0% vs 2%), for DOC (100), DOC (75) and VNB/IFOS, respectively. Differences are not significant
- *QoL*: Reported in another paper (no information provided)

Comments

Methodological comments

- Patients were randomised to each treatment arm. The control arm was VNB/IFOS (dependent on investigator's discretion)
- Baseline characteristics of patients were comparable in all three arms, for age, gender, PS, extent of involvement and histology
- Study was open-label
- ITT analysis was used in survival (primary efficacy analysis). Kaplan–Meier method used for survival and time to progression. Survival curves estimated using Kaplan–Meier method. Overall survival curves compared using the log-rank test. Chi-squared test used to compare other descriptive points. ITT for survival included all randomised patients, whereas ITT for time to progression included only those with a confirmed diagnosis of NSCLC. The ITT for determination of response included patients with confirmed diagnosis of NSCLC who actually received at least one infusion of chemotherapy in the study. The safety population included all patients who received at least one infusion of chemotherapy
- Sample size was based on the assumption that median survival in the group treated with either DOC dose would be approximately 7.5 months, compared with 5 months in the VNB/IFOS control group. Therefore, a sample size of 360 patients (120 per arm) would allow detection of the overall survival advantage in either DOC arm at the alpha level of 5% (one-sided) and 80% power. All statistical tests were performed based on a two-sided error of 5%
- All 373 patients were included in the ITT survival analysis. Three randomised patients (1 in each arm) did not have NSCLC and so were excluded from the time to progression analysis. Twelve patients never received treatment after randomisation (4 patients per arm), and they are excluded from the safety analysis. The 12 non-treated patients and the 3 treated patients without a diagnosis of NSCLC were excluded from the response assessment analysis

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: must have measurable or assessable lesions; adequate bone marrow function (absolute granulocyte count of $\geq 2.0 \times 10^9$ cells/l and platelet count of $\geq 100 \times 10^9$ cells/l); adequate hepatic function (total bilirubin level within normal limits, alkaline phosphatase level ≤ 5 times the upper limit of normal and serum transaminase ≤ 1.5 times the upper limit of normal); adequate renal function (serum creatinine level ≤ 2.0 mg/dl or creatinine clearance ≥ 60 ml/minute); prior radiation was allowed, provided that at least 30 days had elapsed from completion of radiation to study entry; treated brain metastases allowed, provided that they were neurologically stable
- Interventions and outcome measures were clearly described. Toxicity was assessed using the NCI common toxicity criteria. Adverse events not included in that toxicity scale (e.g. fluid retention, hypersensitivity reaction, onychodystrophy and asthenia) were graded as mild (grade 1), moderate (grade 2), severe (grade 3) or life-threatening (grade 4). QoL was assessed using the LCSS
- Prior to study entry, patients underwent baseline evaluation, which included history and physical examination, complete blood count, biochemical profile, ECG, chest X-ray and radiographic imaging of all involved sites of disease by CT scan, bone scan and/or MRI (as clinically appropriate). Weekly assessment included a complete blood count. Prior to each subsequent cycle, history and physical examination, toxicity assessment, complete blood count, biochemical profile and chest X-ray were performed. QoL was assessed at baseline, immediately before each cycle, at end of treatment and at follow-up every 2 months. Tumour responses were assessed radiographically every 2 cycles
- *Conflicts of interest*: Supported by Rhône-Poulenc Rorer

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	3/373 randomised patients (1 in each arm) did not have NSCLC and so were excluded from the time to progression analysis 12/373 patients never received treatment after randomisation (4 patients per arm), and they were excluded from the safety analysis 12/373 non-treated patients and the 3/373 treated patients without a diagnosis of NSCLC were excluded from response assessment analysis

ECOG, Eastern Cooperative Oncology Group; NCI, National Cancer Institute (USA); ECG, electrocardiogram; COSTART, Coding Symbols for Thesaurus of Adverse Reaction Terms; AUC, area under the curve; CT, computed tomography; Mesna, [sodium 2-] mercaptoethane-sulphonate; MRI, magnetic resonance imaging

Appendix 6

Summary of evidence of effectiveness of gemcitabine in lung cancer

Reference and design	Intervention	Participants	Outcome measures
Anderson *et al.*, 2000[14] (UK) Phase II, multicentre, randomised trial Funding: Eli Lilly Jadad score: 3/3	***GEM+BSC arm*** GEM 1000 mg/m^2 (iv) on days 1, 8 and 15 of a 28-day cycle ***BSC arm*** Treated symptomatically; any palliative treatment could be used as clinically indicated, ideally excluding chemotherapy ***Length of treatment*** Maximum of 6 cycles ***Other interventions used*** 74 (49%) of GEM+BSC patients vs 119 (79%) of BSC patients received palliative radiotherapy	*n* = 300 GEM+BSC arm: *n* = 150 BSC arm: *n* = 150 63% of patients were male 60% of patients had locally advanced disease 72% of patients had PS of 60 or 70 ***Characteristics of target population*** Histologically or cytologically proven NSCLC Symptomatic locally advanced or metastatic disease that was not amenable to curative surgery or radiotherapy Median age: GEM+BSC arm, 65 years (range, 37–82 years); BSC arm, 64 years (range, 32–83 years) Previously untreated Clinically measurable disease (uni- or bi-dimensional) PS of 60–90 (Karnofsky scale) Estimated life expectancy ≥ 4 weeks See *General comments* below ***Setting*** Patients allocated to GEM+BSC and BSC arms were treated as outpatients	***Primary end-point*** QoL: Predetermined subset of items (SS14) derived from the EORTC QLQ-C30 and LC13. SS14 included disease-specific items and other most frequently reported symptoms in patients with NSCLC. Evaluated every 4 weeks. Overall QoL (EORTC QLQ-C30 and LC13) also compared ***Secondary end-points*** Survival: Patients followed until time of death (patients alive at the time of analysis were censored at the last date known to be alive) Toxicity: WHO criteria, evaluated every 4 weeks Objective response rate: WHO criteria, evaluated every 4 weeks (for GEM+BSC only) See *General comments* below ***Length of follow-up*** Median follow-up was 25.3 months (range, 1.3–40.3 months)

Results

- *QoL*: Evaluable and unevaluable patients were compared for QoL data. Unevaluable patients had a ≥ 10-point difference in mean and/or median scores for the fatigue and social functioning subscales, appetite loss, constipation and pain (other than chest or shoulder), but were comparable on all other QoL domains. Evaluable patients had greater symptom burden in the cognitive domain
- *QoL (SS14)*: In total, 66% (99) of patients in the GEM+BSC arm and 68% (102) of patients in the BSC arm qualified for SS14 analysis. The percentage change in mean SS14 score from baseline to 2 months was: GEM+BSC, –10.2% (i.e. improvement), BSC, +1% (i.e. deterioration; two sample *t*-test, $p = 0.113$). Sustained (≥ 4 weeks) improvement (≥ 25%) in SS14 score was: GEM+BSC, 22%; BSC, 9% (Pearson's chi-squared test, $p = 0.0014$). SS14 improvement (≥ 25%) in GEM+BSC and BSC arms, respectively: for 2 months, 38% vs 24% (99 and 102 evaluable patients; Pearson's chi-squared test, $p = 0.065$); for 4 months, 44% vs 25% (68 and 61 evaluable patients; Pearson's chi-squared test, $p = 0.015$); for 6 months, 31% vs 22% (36 and 40 evaluable patients; Pearson's chi-squared test, $p = 0.644$). Note: Numbers of patients at 4 and 6 months were insufficient to estimate sustained improvement
- *QoL (overall)*: At 2 months, of the 25 variables analysed, six showed between-treatment differences in improvements that were ≥ 10%: five of the improvements were greater for GEM+BSC (emotional functioning, pain-symptom scale, chest pain, cough and fatigue), whereas one was greater for BSC (dyspnoea). At 2 months, five variables showed between-treatment differences in deterioration that were ≥ 10%: two of the deteriorations were greater for GEM+BSC (role function and hair loss), whereas three were greater for BSC (chest pain, shoulder pain and emotional functioning). At 4 months, six variables showed between-treatment differences in improvements that were ≥ 10%: all six improvements were greater for GEM+BSC (chest pain, shoulder pain, emotional functioning, role domain, social domain and financial impact). Also, at 4 months, four variables showed between-treatment differences in deterioration that were ≥ 10%: one of the deteriorations was greater for GEM+BSC (hair loss), whereas three deteriorations were greater for BSC (social domain, pain-symptom scale and constipation). Improvements in PS (lasting ≥ 4 weeks): GEM+BSC, 20.3%; BSC, 12.3% ($p = 0.073$)
- *Survival*: As of 4 June 1998, 13 patients were still alive. There was no difference in survival between the two arms. Survival curves shown, curves roughly comparable. Median survival: GEM+BSC, 5.7 months (95% CI, 4.6 to 7.6 months); BSC, 5.9 months (95% CI, 5.0 to 7.9 months; log-rank test, $p = 0.84$). Estimated 1-year survival rate: GEM+BSC, 25%; BSC, 22%. Estimated 2-year survival rate: GEM+BSC, 6%; BSC, 7%

continued

***Results* contd**

- *Response*: In the GEM+BSC arm, 15 patients did not have tumour measurements available due to insufficient therapy (11 patients), lack of uni- or bi-dimensional lesions (3 patients) and a diagnosis of mesothelioma (1 patient). Of the 135 patients with at least two assessments of tumour size, 25 patients had objective responses (overall response rate, 18.5%; 95% CI, 13% to 26%)
- *Toxicity*: Incidence of Grade 3–4 toxicity in the GEM+BSC arm was low: neutropenia, 13%; infection, 0.7%; thrombocytopenia, 2%; nausea/vomiting, 9%; lethargy, 6%; rash, 4%; and pulmonary toxicity, 3%. Patient-reported symptoms used to assess chemotherapy toxicity showed that patients in the GEM+BSC arm at 2 months had increased prevalence (compared with the BSC arm) of hair loss (31% vs 6%), ankle swelling (30% vs 11%) and flu-like symptoms (32% vs 15%), but no skin rash (13% vs 16%). RTOG toxicity was low: grade 3 and 4 pharyngeal/oesophageal and skin toxicity was ≤ 2% in each arm; RTOG grade 3 and 4 pulmonary toxicity occurred in 4% of BSC patients who received radiotherapy, but there was none in the patients in the GEM+BSC arm who subsequently received radiotherapy

Comments

Methodological comments

- Computer-generated randomisation was performed centrally by telephone. Patients were stratified for 25 treatment centres, PS (Karnofsky PS 80–90 vs 60–70) and disease extent (locoregional vs metastatic)
- Patients were well matched for pretreatment characteristics: age, gender, Karnofsky PS and stage
- Blinding was not specified
- Results were presented on an ITT basis. The percentage change in the mean score of the SS14 items in each randomised group, from baseline to 2 months, was compared using a two sample *t*-test. The difference in sustained symptom improvement rates was assessed using Pearson's chi-squared test. Overall survival curves were produced using the Kaplan–Meier method and were compared using the log-rank test. Baseline QoL forms were included only if completed on or before randomisation, but acceptable time windows of ± 1 week were permitted around QoL assessment points of 2, 4 and 6 months. In total, 67% of patients randomised were evaluable for analysis of QoL data with respect to the primary end-point
- The study was designed to recruit 300 patients, with 150 in each arm. The trial was designed so that the sample size of 150 patients per arm would provide 90% power to detect a difference of 0.4 SD at the 5% significance level
- One patient was found to have mesothelioma, but results were presented on an ITT basis. In the GEM+BSC arm, 51 patients did not qualify for primary QoL analysis: 10 did not complete QoL forms at specified times, 22 died by 2 months, and 19 had QoL forms with data missing. In the BSC arm, 49 patients did not qualify for primary QoL analysis: 16 did not complete QoL forms at specified time points, 22 had died by 2 months, and 19 had QoL forms with missing data

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: symptomatic locally advanced or metastatic disease that was not amenable to curative surgery or radiotherapy, no urgent radiotherapy required, no brain metastases, adequate bone marrow reserve (leucocyte count < $3.5 \times 10^9/l^{-1}$, platelets < $100 \times 10^9/l^{-1}$ and Hb < 100 g/l^{-1}) and adequate liver function (bilirubin > 3 times above normal range; ALT or AST > 3 times normal)
- Interventions were clearly defined. Outcome measures were clearly defined. QoL end-points used to assess change in symptoms were: the percentage change in mean SS14 score from randomisation to 2 months; and the proportion of patients with sustained improvement of SS14 score at 2 months, defined as ≥ 25% reduced from baseline sustained from month 1 to month 2, and/or from month 2 to month 3. Multidimensional QoL parameters: patient-assessed QoL using all the subscales and symptom items on the QoL measures. Changes from baseline to 2, 4 and 6 months were calculated in terms of the proportion of patients who improved or deteriorated. Toxicity was assessed based on WHO criteria. For objective tumour response rate among patients receiving GEM+BSC, tumour response was defined according to WHO criteria. Three additional symptoms were included to assess possible GEM side-effects (skin rash/itchiness, ankle swelling, flu-like symptoms). Patients completed EORTC QLQ-C30 and LC13 questionnaires every 4 weeks, prior to their clinical assessment
- For GEM+BSC patients, FBC was performed weekly during first cycle and every 2 weeks thereafter. BSC patients were seen in clinic every 4 weeks
- *Conflicts of interest:* Supported by Eli Lilly and Company

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	Not possible due to comparison with BSC
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	One patient was found to have mesothelioma, but results were presented on an ITT basis 67% of patients randomised were evaluable for analysis of QoL data with respect to the primary end-point: 99 patients in GEM+BSC arm and 102 patients in BSC alone arm In GEM+BSC arm, 51 patients did not qualify for primary QoL analysis: 10 did not complete QoL forms at specified times, 22 died by 2 months, 19 had QoL forms with data missing In BSC arm, 49 patients did not qualify for primary QoL analysis: 16 did not complete QoL forms at specified time points, 22 had died by 2 months, 19 had QoL forms with missing data

Reference and design	Intervention	Participants	Outcome measures
Bokkel-Huinink *et al.*, 1999[24] Update of interim analysis provided by Manegold *et al.*, 1997[25] (International) Phase II, multicentre, open-label, randomised study Funding: Eli Lilly Jadad score: 2/5	***GEM arm*** GEM 1000 mg/m^2 (iv) on days 1, 8 and 15 of 28-day cycle ***CDDP+VP-16 arm*** CDDP 100 mg/m^2 (iv) on day 1 and VP-16 100 mg/m^2 (iv) on day 1 (after CDDP), 2 and 3 of 28-day cycle ***Length of treatment*** Maximum number of 6 cycles, unless evidence of disease progression ***Other interventions used*** Prednisone 10 mg/m^2, prophylactic 5-HT$_3$ antagonist antiemetics (except cycle 1 in GEM arm), dexamethasone 20 mg or equivalent (CDDP+VP-16 arm only). Palliative local radiotherapy (provided other measurable sites were being assessed). Growth factors for prolonged myelo-suppression	n = 147 patients GEM arm: n = 72 CDDP+VP-16 arm: n = 75 78% of patients were male 86% of patients had PS of 0 or 1 76% of patients had stage IV NSCLC 78% of patients had adenocarcinoma or squamous cell carcinoma ***Characteristics of target population*** Histological/cytological diagnosis of NSCLC stage IIIA (if inoperable), IIIB or IV Aged ≥ 18 years Median age: GEM arm, 59 years (range, 32–80 years); CDDP+VP-16 arm, 59 years (range, 33–78 years) PS ≤ 2 (Zubrod scale) Clinically measurable disease No prior or concurrent chemotherapy No prior radiotherapy (except where irradiated area was not only source of measurable disease) See *General comments* below ***Setting*** Not specified	***Primary end-points*** Objective tumour response: WHO response criteria. Evaluation at end of each treatment cycle Tolerability: WHO grading for toxic effects. Evaluation at end of each treatment cycle ***Secondary end-points*** QoL: EORTC QLQ-C30-LC13, completed at baseline and end of each therapy cycle Disease-related symptoms: Assessed at baseline and every therapy cycle Time-to-event efficacy parameters: Overall patient survival, duration of response, time to progressive disease, and time and duration of tumour response. Assessed at baseline and every other therapy cycle See *General comments* below ***Length of follow-up*** 24 months after last patient enrolled

Results

- *Haematological toxicity*: Grades 3 and 4: anaemia (7% and 10% vs 0% and 0%), neutropenia (7% and 9% vs 1% and 36%), thrombocytopenia (1% and 10% vs 0% and 10%) for GEM and CDDP+VP-16 arms, respectively. More patients in the CDDP+VP-16 arm experienced grade 3 and 4 neutropenia (p = 0.0000003). Grade 3 and 4 thrombocytopenia was significantly higher for CDDP+VP-16 arm (p = 0.003)
- *Clinical toxicity*: Grades 3 and 4: hair loss (0% and 61% vs 0% and 1%), nausea/vomiting (11% and 26% vs 0% and 4%), infection (4% and 4% vs 0% and 4%), diarrhoea (0% and 3% vs 0% and 0%), pulmonary (4% and 4% vs 6% and 0%) for GEM and CDDP+VP-16 arms, respectively. CDDP+VP-16 arm had notably higher incidence and severity of clinical toxicity
- *QoL*: Of 147 individuals enrolled, 125 (85%) met the criteria of completing a baseline and at least one postbaseline QoL questionnaire: GEM arm, 64 patients; CDDP+VP-16 arm, 61 patients. Median number of questionnaires answered by both arms: 3
- *Functional scales (physical, role, cognitive, emotional and social)*: No statistically significant difference between the two treatment groups in change from mean baseline scores over 6-cycle treatment period (p > 0.05). No statistically significant difference in change from mean baseline score for global QoL either between or within treatment arms (p > 0.05)
- *Symptom scales*: Statistically significant differences in change from mean baseline scores (p < 0.05) were not indicative of trends across treatment period, except for hair loss, which worsened for the CDDP+VP-16 arm after all cycles
- *Mean EORTC QLQ-C30-LC13 scores*: Reported for each arm at baseline and at month 2 for select symptoms
- *Survival curve (Kaplan–Meier)*: No difference in overall survival between both GEM and CDDP+VP-16 curves
- *One-year survival probability*: GEM, 26%; CDDP+VP-16, 24%
- *Median survival time (months)*: GEM, 6.6 (95% CI, 4.9 to 7.3); CDDP+VP-16, 7.6 (95% CI, 5.4 to 9.3)
- *Median time to treatment failure (months)*: GEM, 2.8 (95% CI, 2.0 to 3.7); CDDP+VP-16, 2.8 (95% CI, 2.0 to 3.6)
- *Median time to progressive disease (months)*: GEM, 3.0 (95% CI, 2.2 to 3.9); CDDP+VP-16, 3.2 (95% CI, 2.1 to 4.8)
- *Median duration of tumour response (months)*: GEM, 6.5 (95% CI, 3.8 to 9.8); CDDP+VP-16, 5.8 (95% CI, 4.8 to 7.2)
- *Median time to tumour response (months)*: GEM, 1.9 (95% CI, 1.1 to 3.0); CDDP+VP-16, 1.9 (95% CI, 1.9 to 2.9)
- *Efficacy (protocol-qualified)*: Complete response: GEM, 0%; CDDP+VP-16, 0%. Partial response: GEM, 17.9%; CDDP+VP-16, 15.3%. Stable disease: GEM, 41.8%; CDDP+VP-16, 51.4%. Progressive disease: GEM, 28.4%; CDDP+VP-16, 19.4%. Not evaluated: GEM, 11.9%; CDDP+VP-16, 13.9%
- No statistically significant difference between the two treatment groups for any of the parameters (p > 0.05)
- Tumour response showed no statistically significant difference between the two therapy groups (log-rank test, p = 0.37; Wilcoxon test, p = 0.97)
- Duration of tumour response showed no statistically significant difference between the two groups (log-rank test, p = 0.76; Wilcoxon test, p = 0.67)

continued

Comments

Methodological comments

- Patients were randomised, but no indication of method of randomisation. Open-label trial. Not placebo-controlled. Stratified according to stage, PS and investigator centre
- Patient demographics and baseline disease characteristics were well matched, although slightly more patients in GEM arm (16.7% vs 9.3%) with PS of 2 (Zubrod scale)
- Intended sample size: 140 qualified patients. Number considered enough to allow toxicity and tolerability profile of GEM to be compared with CDDP+VP-16 and with toxicity and tolerability profiles of prior GEM studies. The study was not powered to detect differences in objective tumour response rates (because it was assumed that response rates of the two treatment arms would be similar). Therefore, efficacy data from this study should be viewed with caution regarding significant effects (with the exception of toxicity data)
- ITT analysis was not used ("protocol-qualified" patients used for efficacy analysis). All patients randomised to treatment and meeting the eligibility criteria (protocol-qualified) were used for objective tumour response rate analysis. All patients who received at least one dose of GEM or CDDP+VP-16 were evaluated for safety. For survival, data for patients alive at the cut-off date were right-censored prior to analysis. For patients lost to follow-up before the cut-off date, the last date the patient was known to be alive was used. Protocol-qualified patients (n = 139; GEM, n = 67; CDDP+VP-16, n = 72) and all randomised patients (n = 147) were analysed with time-to-event parameters. Time-to-event data in this paper report only protocol-qualified patients. All patients completing baseline and at least one postbaseline QoL questionnaire were included in QoL analysis (n = 125; GEM, n = 64; CDDP+VP-16, n = 61)
- Five patients in GEM arm (2 due to rapid disease progression, 3 due to no bi-dimensionally measurable tumour lesions at entry) and 3 in CDDP+VP-16 arm (2 due to rapid disease progression, 1 due to no bi-dimensionally measurable tumour lesions at entry) were not qualified for efficacy analysis
- Concern exists regarding the use of 'objective tumour response rates' as a primary outcome, when study was not powered to detect differences
- QoL analysis represents only those patients who remained in the study; analysis of data was limited by the high dropout rate, which was similar between the two arms

General comments

- Eligibility and exclusion criteria were clearly defined. Additional inclusion criteria included: adequate bone marrow reserve (WBC, platelets, Hb, haematocrit), no symptomatic nervous system metastases, adequate liver and renal function, no concomitant treatment with nephrotoxic antibiotics)
- Study outcomes were defined (unclear as to why response rate was used as primary end-point). Interventions were clearly defined
- Baseline disease status was assessed no more than 1 week before enrolment (medical history, tumour palpation, evaluation of PS, chest X-ray and physical examination performed). Before each therapy cycle, body weight and PS were determined. Before every other cycle, chest X-ray and radiological imaging studies were performed. Treatment responders had chest X-rays and radiological imaging 4 weeks after initial classification
- Concern exists that there was a gross lack of reporting regarding tolerance or toxicity parameters/outcomes
- *Conflicts of interest*: Study funded by Eli Lilly

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	8/147 patients (5%) dropped out of efficacy analysis: 5/72 patients (7%) in GEM arm (2 because of rapid disease progression, 3 because of no bi-dimensionally measurable tumour lesion at entry) 3/75 patients (4%) in CDDP+VP-16 arm (2 because of rapid disease progression, 1 because of no bi-dimensionally measurable tumour lesion at entry)

Reference and design	Intervention	Participants	Outcome measures
Manegold *et al.*, 1997[25] (Europe) Phase II, open-label, multicentre, randomised trial Funding: Not specified Jadad score: 2/5	***GEM arm*** GEM 1000 mg/m^2 (iv) on days 1, 2 and 15 of a 28-day cycle ***CDDP+VP-16 arm*** CDDP 100 mg/m^2 (iv) on day 1, with VP-16 100 mg/m^2 (iv) on days 1 (following CDDP), 2 and 3 of a 28-day cycle ***Length of treatment*** Maximum of 6 cycles or until disease progression ***Other interventions used*** Prednisone, prophylactic 5-HT$_3$ receptor blocking agents and steroids (dexamethasone or equivalent), prophylactic 5-HT$_3$ antiemetics (though not in cycle 1), palliative radiotherapy (if an existing lesion became painful without being accompanied by other objective changes indicating disease progression), provided other measurable sites were being assessed	*n* = 146 GEM arm: *n* = 71 CDDP+VP-16 arm: *n* = 75 61% of patients were male 87% of patients had PS of 0 or 1 75% of patients had stage IV disease 47% of patients had adenocarcinoma histotype ***Characteristics of target population*** Histological diagnosis of NSCLC Stage IIIA (if inoperable), IIIB or IV, according to AJCC criteria Aged ≥ 18 years Median age: GEM, 59 years (range, 32–80 years); CDDP+VP-16, 59 years (range, 33–78 years) PS of ≤ 2 (Zubrod scale) No prior chemotherapy No prior radiation therapy, except when the irradiated area was not the only source of measurable disease See *General comments* below ***Setting*** Not specified	***Primary end-point*** Objective response: WHO criteria. Time of evaluation not specified ***Secondary end-points*** Time to treatment failure: Measured from the time of first dose of study drug(s) to discontinuation Survival: Calculation not defined Toxicity: WHO criteria and time to evaluation not specified ***Length of follow-up*** Date of early analysis was 1 July 1996, after randomisation occurred between June and December 1995

Results

- *Response*: In total, 93% of GEM patients and 96% of CDDP+VP-16 patients were evaluable for efficacy analysis. No complete response was observed in either therapy arm. Partial response: 12 patients (18.2%; 95% CI, 9.8% to 30%) vs 11 patients (15.3%; 95% CI, 7.9% to 25.7%); stable disease, 30 patients (45.5%) vs 35 patients (48.6%), for GEM and CDDP+VP-16, respectively. Overall response was the same as partial response. No significant differences between arms
- *Median time to progression*: GEM, 4.2 months (95% CI, 2.9 to 5.6 months); CDDP+VP-16, 4.9 months (95% CI, 3.2 to 5.8 months; $p > 0.90$). Patients progression-free at time of analysis: GEM, 35%; CDDP+VP-16, 38%
- *Survival*: Randomisation between June and December 1995. Data collected as of 1 July 1996. Patients still alive at time of analysis: GEM, 50%; CDDP+VP-16, 46%. Median survival: GEM, 6.6 months (95% CI, 4.9 to 7.1 months); CDDP+VP-16, 7.6 months (95% CI, 5.6 to 9.6 months; $p > 0.90$)
- *Haematological toxicity*: Grade 3: neutrophils (6% vs 3%), Hb (4% vs 3%), platelets (2% vs 0%), for GEM and CDDP+VP-16 arms, respectively. Grade 4: neutrophils (2% vs 12%), for GEM and CDDP+VP-16 arms, respectively. No significance levels provided
- *Non-haematological toxicity*: Grade 3: alopecia (3% vs 60%), nausea/vomiting (11% vs 25%), neurohearing (0% vs 6%), dyspnoea (4% vs 4%), for GEM and CDDP+VP-16 arms, respectively. Grade 4: alopecia (0% vs 2%), nausea/vomiting (0% vs 4%), dyspnoea (6% vs 0%), for GEM and CDDP+VP-16 arms, respectively. No significance levels provided

Comments

Methodological comments

- Patients were randomised by open-label and were stratified by stage, PS and investigator centre using the Pocock and Simon algorithm
- Patients' prognostic factors were well balanced across both treatment arms, with slightly more cases of PS 2 in the GEM arm and of stage IIIA in the CDDP+VP-16 arm
- Study was open-label
- All patients receiving at least one dose of GEM or CDDP+VP-16 were evaluated for safety. All patients randomised to treatment and meeting the eligibility criteria were considered qualified for objective tumour response assessment and for analysis of the time-to-event parameters. An ITT analysis of all randomised patients was made
- Sample size/statistical power not indicated
- Five (7%) of GEM patients were not evaluable for efficacy analysis (2 with rapid disease progression, 3 with no bi-dimensionally measurable disease). Three (4%) of CDDP+VP-16 patients were not evaluable for efficacy analysis (2 with rapid disease progression, 1 with no bi-dimensionally measurable disease)

continued

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate bone marrow reserve (WBC ≥ 3.5 x 10^9/l, platelets ≥ 100 x 10^9/l, Hb ≥ 100 g/l), no CNS metastases, bilirubin > 1.5 times normal, prothrombin time or activated partial thromboplastin time > 1.5 times control, ALT or AST > 3 times normal (or up to 5 times normal in patients with known liver metastases), serum calcium levels above normal limits, no serious concomitant systemic disorders, no concomitant treatment with nephrotoxic antibiotics
- Interventions were clearly defined. Outcome measures were not defined clearly (no indication of evaluation time for toxicity or calculation of survival time)
- Weekly absolute granulocyte and platelet counts were taken on day of therapy. Serum creatinine was obtained prior to each cycle
- *Conflicts of interest*: Not specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0 (open-label)
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	5 (7%) of GEM patients were not evaluable for efficacy analysis (2 with rapid disease progression, 3 with no bi-dimensionally measurable disease) 3 (4%) of CDDP+VP-16 patients were not evaluable for efficacy analysis (2 with rapid disease progression, 1 with no bi-dimensionally measurable disease)

Reference and design	Intervention	Participants	Outcome measures
Cardenal *et al*., 1999[26] (Spain) Phase III, multicentre, randomised trial Funding: Eli Lilly Jadad score: 2/5	***GEM arm*** GEM 1250 mg/m^2 (iv) on days 1 and 8 of 21-day cycle, with CDDP 100 mg/m^2 (iv) on day 1 of 21-day cycle ***VP-16 arm*** VP-16 100 mg/m^2 (iv) on days 1, 2 and 3 of 21-day cycle, with CDDP 100 mg/m^2 (iv) on day 1 of 21-day cycle ***Length of treatment*** Maximum of 6 cycles ***Other interventions used*** 5% glucose+saline before GEM or VP-16. Supportive care (e.g. blood-product transfusions, antibiotics, antiemetics, analgesics). Palliative radiation therapy was allowed for a previous painful lesion, as long as the patient was not qualified as having progressive disease and the irradiated lesion was not the only measurable lesion (21 patients in GEM arm and 14 patients in VP-16 arm received radiation therapy)	*n* = 135 GEM arm: *n* = 69 VP-16 arm: *n* = 66 93% of patients were male 85% of patients had good/excellent PS (80–100) 50% of patients had stage IIIB or unresectable local recurrence ***Characteristics of target population*** Histological/cytological diagnosis of NSCLC stage IIIB or IV Aged ≥ 18 years Median age: GEM arm, 59 years (range, 33–76 years); VP-16 arm, 58 years (range, 35–75 years) PS ≥ 60 (Karnofsky criteria) Life expectancy ≥ 12 weeks No prior chemotherapy 1+ measurable lesion At least 3 weeks elapsed since prior radiotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response rate: Common response criteria. Indicated lesions were measured no more than 3 weeks before randomisation and just before every other cycle ***Secondary end-points*** Toxicity: WHO criteria (ototoxicity of CDDP graded by common toxicity criteria). Recorded before every scheduled treatment (regardless of delays) QoL: EORTC QLQ-C30-LC13 completed at baseline and end of every cycle Time to disease progression: Defined as the time from the date of randomisation to the date the patient was assessed as having progressive disease Survival time: Defined as the interval between the date of randomisation and the date of death ***Length of follow-up*** Minimum follow-up time was 16 months See *General comments* below

Results

- *Response*: No patients achieved complete response (ITT used). In total, 28 (40.6%) of patients in the GEM arm (95% CI, 29% to 53%) and 14 (21.9%) of patients in the VP-16 arm (95% CI, 13% to 34%) achieved a partial response (p = 0.02). Response rates were stratified in terms of age, sex, PS and disease stage. Response rates were 36% for stage IIIB and 44% for stage IV patients in the GEM arm, and 26% for stage IIIB and 17% for stage IV patients in the VP-16 arm. The majority of responders had a PS of 80–100
- *QoL:* In total, 68 patients in GEM arm and 63 patients in VP-16 arm completed at least one QoL questionnaire. Median number of questionnaires completed was 5.5 in GEM arm and 4 in VP-16 arm, with an overall on-study compliance rate of 88.1% and 84.5%, respectively. No clinically significant differences in change from baseline within treatment arm or between treatment arms in functional domains or global QoL were observed. Both groups saw significant improvement in pain, insomnia, cough, haemoptysis, chest pain and shoulder pain. No improvement in dyspnoea and fatigue was evident for either arm. Peripheral neuropathy did not worsen in either arm. Both arms had significant worsening of nausea and alopecia. The only statistically significant difference between treatment arms in change from baseline was for alopecia, which was worse for VP-16 arm (significance levels not provided)
- *Disease progression*: Time to disease progression curves shown; GEM arm curve is above the VP-16 arm curve. Median time to progression for GEM patient was 6.9 months (95% CI, 5 to 8.1 months), compared with 4.3 months (95% CI, 3.5 to 4.7 months) for VP-16 patients (log-rank test, p = 0.01). Probability of tumour response lasting at least 6 months was estimated to be 79% for GEM arm and 57% for VP-16 arm. Proportional hazards analysis suggested no prognostic factor (age, sex, disease stage, PS, prior radiation therapy, liver metastases, time since diagnosis) was significantly related to survival time or time to disease progression
- *Survival*: Survival curves shown; GEM arm survival curve is above the VP-16 arm. GEM patients had an estimated median survival time of 8.7 months (95% CI, 7.7 to 10.2 months), and VP-16 patients had an estimated median survival time of 7.2 months (95% CI, 6.1 to 9.8 months). Difference between the two curves was statistically significant (p = 0.02). One-year survival probability estimated as 32% for GEM arm and 26% for VP-16 arm. Difference between two arms was not statistically significant (log-rank test, p = 0.19). The censoring rates (at time of analysis, 15 January 1998 – randomisation between July 1995 and June 1996) for the survival curves were 16% for GEM arm and 11% for VP-16 arm
- *Haematological toxicity*: Main toxicity was myelosuppression. Grade 3 and 4 neutropenia and febrile neutropenia were more pronounced in the VP-16 arm. Grade 4 neutropenia was twice as frequent in VP-16 arm in comparison with GEM arm (p = 0.0009). Grades 3 and 4: anaemia (22% and 0% vs 13% and 2%), neutropenia (36% and 28% vs 20% and 56%), thrombocytopenia (39% and 16% vs 8% and 5%), for GEM and VP-16 arms, respectively. Incidence of febrile neutropenia (7% vs 12%), PRBC transfusion (29% vs 21%), platelet transfusion (3% vs 8%), toxic deaths (1% vs 0%), for GEM and VP-16 arms, respectively

continued

***Results* contd**

- *Non-haematological toxicity*: Grades 3 and 4: nausea/vomiting (35% and 4% vs 19% and 7%), haemorrhage (0% and 3% vs 0% and 3%), fever (0% in both arms), infection (3% and 1% vs 8% and 0%), dyspnoea (0% and 1% vs 0% and 0%), alopecia (13% and 0% vs 51% and 0%), paresthesias (0% and 0% vs 2% and 0%), for GEM and VP-16 arms, respectively

Comments

Methodological comments

- In total, 135 patients were randomised. Not placebo-controlled. Two patients did not meet eligibility criteria for efficacy analysis, one due to non-measurable disease, and one had less than 1 cycle of therapy received
- Baseline characteristics were well balanced in both arms of trial. Chi-squared test for homogeneity was carried out on prognostic factors
- There was no indication of blinding
- ITT analysis was carried out. Cox proportional hazards exploratory analysis was performed using potential prognostic factors of age, sex, disease stage, PS, prior radiation therapy, liver metastases and time since diagnosis as co-variates. Logistic regression was carried out to indicate that the single predictor of response was treatment. Wilcoxon and log-rank tests were carried out to determine that GEM+CDDP provided a significantly longer time to disease progression. Fisher's exact test was used to assess response rates
- Sample size was estimated to be 62 patients per arm to provide a power of 0.80 to detect a 25% difference in response rates between the two groups at the 5% level
- Two individuals dropped out, one due to non-measurable disease and one due to having less than 1 cycle of therapy

General comments

- Overall number of patients alive at end of study was not reported
- Eligibility and exclusion criteria were clearly defined. Additional inclusion criteria included: no prior malignancy, except basal cell carcinoma of the skin or carcinoma *in situ* of the cervix; adequate bone marrow, renal and hepatic function; patients who had received previous radiotherapy were accepted if their assessable disease was outside the radiation portal; haematopoietic growth factors permitted in the case of prolonged neutropenia
- Interventions and outcomes were clearly defined
- Complete blood counts were performed every week on patients in GEM arm, and on days 1 and 15 of 21-day cycle on patients in the VP-16 arm. Blood chemistry, urine analysis and a physical examination were repeated at the beginning of every cycle. Chest X-ray was repeated before every other cycle. Complete response was defined as complete disappearance of all objective disease. Partial response was defined as a ≥ 50% reduction in the size of all measurable tumour areas from baseline. Complete and partial responses were confirmed by second evaluation at least 4 weeks later. All responses were peer-reviewed by two independent radiologists
- *Conflicts of interest*: Funded by Eli Lilly

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	2/135 patients dropped out: 1/66 patients in VP-16 (with CDDP) arm due to death 1/66 patients in VP-16 (with CDDP) arm due to non-measurable disease

Reference and design	Intervention	Participants	Outcome measures
Crino *et al.*, 1999[27] (Italy) Phase III, multicentre randomised trial Funding: No information provided Jadad score: 2/5	***GEM+CDDP arm*** GEM 1000 mg/m^2 on days 1, 8 and 15 of 28-day cycle, with CDDP 100 mg/m^2 on day 2 of 28-day cycle ***TriCOMB arm*** MIC 6 mg/m^2 on day 1, IFOS 3000 mg/m^2 on day 1 and CDDP 100 mg/m^2 on day 2 of 28-day cycle ***Length of treatment*** 6 cycles ***Other interventions used*** Mesna at dose equivalent to 20% of IFOS dose at time of drug administration and 4–8 hours after (TriCOMB arm only). Antiemetics (5-HT antagonists plus corticosteroids) and other supportive therapy	*n* = 307 patients GEM+CDDP arm: *n* = 155 TriCOMB arm: *n* = 152 85% of patients were male 94% of patients had PS of 0 or 1 79% of patients had stage IV disease 17% of patients in GEM+CDDP arm and 15% in TriCOMB arm had brain metastases at study entry ***Characteristics of target population*** Aged between 18 and 75 years Median age: GEM+CDDP arm, 62 years (range, 28–76 years); TriCOMB arm, 60 years (range, 25–75 years) Histologically/cytologically diagnosed NSCLC; stage IIIB (limited to T4 for pleural effusion and N3 for supraclavicular lymph nodes) and IV PS ≤ 2 (Zubrod scale) No prior chemotherapy, immunotherapy or radiotherapy Life expectancy ≥ 12 weeks 1+ uni-dimensionally (though prefer bi-dimensionally) measurable lesion See *General comments* below ***Setting*** Not specified	***Primary end-point*** QoL: Self-administered EORTC QLQ-C30-LC13 and physician assessment of disease-related symptoms. Evaluated before each therapy cycle ***Secondary end-points*** Response rates: WHO criteria evaluated before every therapy cycle Survival: Defined as the interval between the date of randomisation and the date of death Toxicity: WHO grading evaluated before every therapy cycle See *General comments* below ***Length of follow-up*** Not specified

Results

- *QoL*: Selected EORTC QLQ-C30-LC13 scores for patients who completed baseline and another questionnaire at cycle 2 are reported. Median number of questionnaires completed was 5 in GEM+CDDP arm and 4 in TriCOMB arm. In total, 155 (100%) patients in GEM+CDDP arm and 151 (99%) in TriCOMB arm completed at least one QoL questionnaire. Overall, there were no differences in changes in QoL between the two arms. The only differences between the two treatment arms for change from baseline were a worsening of alopecia in the TriCOMB arm and a greater improvement in chest pain in the GEM+CDDP arm ($p < 0.05$). Global QoL did not change significantly in either treatment arm. Both treatment arms noted a moderate decrease of physical functioning, also evidenced by worsening of fatigue and nausea/vomiting. Both arms noted improvement of pain, insomnia and cough
- *Response rate*: Complete response, 1% vs 1%; partial response, 37% vs 25%; stable disease, 40% vs 45%; progressive disease, 22% vs 29%, for GEM+CDDP and TriCOMB arms, respectively. Overall response was significantly higher ($p = 0.029$) in GEM+CDDP, 38% (95% CI, 31% to 46%), than TriCOMB arm, 26% (95% CI, 19% to 33%). Median duration of response was 8.7 months for GEM+CDDP arm and 8.2 months for TriCOMB arm. No difference in response rate between the two arms for patients with stage IIIB disease (GEM+CDDP, 41%; TriCOMB, 37%). Response rate for stage IV disease was superior ($p = 0.02$) in GEM+CDDP arm (37%) vs TriCOMB arm (23%)
- *Survival*: In total, 44 (28%) patients in the GEM+CDDP arm and 38 (25%) in the TriCOMB arm were alive at the end of the study. Survival curves shown; two curves comparable. No significant difference (log-rank test, $p = 0.8771$; Wilcoxon, $p = 0.3393$) in overall median survival time between GEM+CDDP (8.6 months) and TriCOMB (9.6 months) arms. One-year survival rate: GEM/CDDP arm, 33%; TriCOMB arm, 34%. No differences for median time to progression (log-rank test, $p = 0.6938$; Wilcoxon, $p = 0.7794$) between the two arms (GEM+CDDP, 5 months; TriCOMB, 4.8 months) or median time to treatment failure (GEM+CDDP, 4 months; TriCOMB, 3.7 months)
- *Haematological toxicity*: Myelosuppression was main toxicity in both arms. Grades 3 and 4: neutropenia (29% and 16% vs 17% and 17%), anaemia (24% and 7% vs 20% and 5%), thrombocytopenia: (26% and 38% vs 16% and 12%), for GEM+CDDP and TriCOMB arms, respectively. Grade 4 thrombocytopenia was significantly more frequent and severe in the GEM+CDDP arm ($p < 0.001$)
- *Non-haematological toxicity*: Grades 3 and 4: severe nausea/vomiting (17% and 1% vs 31% and 21%), alopecia (12% and 0% vs 39% and 0%), for GEM+CDDP arm and TriCOMB arm, respectively. Alopecia was more frequent in TriCOMB arm ($p < 0.001$). Grade 3 peripheral neuropathy was rare (GEM+CDDP, 1 individual; TriCOMB, 2 individuals). Grade 4 peripheral neuropathy was not observed. Moderate and severe dyspnoea (grades 3 and 4): GEM+CDDP arm, 6 patients (5 patients with grade 3 and 1 patient with grade 4); TriCOMB arm, 3 patients (all grade 3)

continued

Comments

Methodological comments

- Randomisation was performed centrally. Not placebo-controlled. Patients were stratified by extent of disease, PS and investigation site
- Prognostic factors of age, gender, PS, stage of disease and histology were equally distributed in both arms
- Blinding was not specified
- ITT analysis was used for efficacy and toxicity data. ITT was not used for QoL due to missing data
- Sample size was based on QoL, although suitable experience and literature on the scale were not available when the study was planned. Sample size of 150 individuals per arm gave the trial at least a 90% chance of detecting a difference between treatment groups in change from baseline to 2 months in a subset of 14 symptom items from the EORTC QLQ-C30-LC13. Assumes minimum clinically meaningful treatment group difference (d) is related to the observed between-group variance (σ). The ratio of $d/\sigma = 0.4$ is recommended, and the sample size was sufficient to detect treatment differences with this ratio or higher. Sample size was not calculated using objective tumour response, which is more conventional. However, a sample size of 150 patents per arm allowed the detection of a difference, with at least 80% power, in the tumour response rate of 18%. Both calculations assumed a two-sided significance level of 0.05 and an estimated non-assessability rate for efficacy analysis of 13%
- In total, 337 patients were initially recruited. Of these, 307 patients qualified for randomisation: 155 in GEM+CDDP arm and 152 in TriCOMB arm. One patient from TriCOMB arm refused treatment after randomisation. Thirty patients failed to qualify for randomisation due to misdiagnosis, patient decision or poor health. Eight individuals in each arm died early. Nearly all deaths were unrelated to study medications, except for three: two in GEM+CDDP arm (one case of acute renal failure after CDDP administration and one case of heart failure) and one in TriCOMB arm (sudden death). In the remaining 13 cases, nine deaths were due to rapid clinical progression of the neoplastic disease (six in TriCOMB arm and three in GEM+CDDP arm), and four (all in GEM+CDDP arm) were due to heart arrest (two cases), sudden death (one case) or respiratory failure (one case). Concern exists that the number of deaths (16) referred to in the text is different from the number (24) presented in *Table 2* of the article by Crino and co-workers[27]

General comments

- Exclusion and eligibility criteria were clearly described. Additional inclusion criteria included: no second primary malignancy, except *in situ* carcinoma of the cervix or adequately controlled basal cell carcinoma of skin; no active infection; adequate bone marrow reserve (WBC, platelet, Hb and haematocrit); normal liver and renal function; patients with brain metastases included in the study if they did not require emergency therapy
- End-points of the trial were defined, although concern exists that time of evaluation was not clearly specified for either response rate or toxicity
- Baseline evaluation, QoL assessment, complete history and physical examination, FBC and serum chemistry analysis, urinalysis, ECG, chest X-rays in the posteroanterior and lateral views and CT scans for tumour measurement were performed within 4 weeks before study entry. Baseline evaluations were repeated after 2 cycles and, if response occurred, then again 4 weeks later. FBC was repeated on day 1 in both arms and weekly in GEM+CDDP arm before each GEM administration
- *Conflicts of interest*: Not specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out? (Caution should be taken when interpreting the TriCOMB data because some inconsistencies are evident.)	1/152 patients (0.7%) in TriCOMB arm refused treatment after randomisation 8/155 patients (6%) in GEM+CDDP arm died early: 2/8 deaths (22%) related to study medications (1/2 [50%] due to acute renal failure after CDDP administration, 1/2 [50%] due to heart failure), 3/8 (37.5%) due to rapid clinical progression of the neoplastic disease, 2/8 (25%) due to heart arrest, 1/8 (12.5%) sudden death, 1/8 (12.5%) due to respiratory arrest 8/152 patients (5%) in TriCOMB arm died early: 1/8 deaths (12.5%) related to study medications (sudden death), 6/8 (75%) due to rapid clinical progression of the neoplastic disease

Reference and design	Intervention	Participants	Outcome measures
Perng *et al*., 1997[28] (China) Phase II, randomised trial Funding: No information provided Jadad score: 3/5	***GEM arm*** GEM 1250 mg/m^2 (iv) on days 1, 8 and 15 of each 28-day cycle ***COMB arm*** CDDP 80 mg/m^2 (iv) on day 1 and VP-16 80 mg/m^2 (iv) on days 1, 2 and 3 of each 28-day cycle ***Length of treatment*** Responding patients and those with stable disease continued on treatment until disease progression or after six courses of treatment ***Other interventions used*** Dexamethasone, metoclopramide, granisetron, lorazepam and saline	*n* = 53 patients GEM arm: *n* = 27 COMB arm: *n* = 26 64% of patients were male 74% of patients had stage IV NSCLC 79% of patients had PS of 0 or 1 68% of patients had adenocarcinoma ***Characteristics of target population*** Histological or cytological diagnosis Inoperable stage III (A or B) or IV NSCLC Aged 18–75 years Median age: GEM arm, 63 years (range, 36–75 years); COMB arm, 60 years (range, 35–75 years) No prior chemotherapy, immunotherapy or radiotherapy PS ≤ 2 (Zubrod scale) Clinically measurable lesions Estimated life expectancy of at least 12 weeks See *General comments* below ***Setting*** GEM administered to outpatients COMB administered to inpatients	***Primary end-point*** Objective response rate: WHO criteria. Evaluated after the first course of chemotherapy and every two courses thereafter ***Secondary end-points*** Survival time: Measured from the date of initial administration of the study drug until the date of death Time to disease progression: Measured from the date of initial administration of the study drug until the date of death Toxicity: WHO criteria, FBC and urinalysis repeated every week in GEM arm and every 2 weeks in COMB arm; serum biochemistry before every therapy cycle See *General comments* below ***Length of follow-up*** Not stated

Results

- *Response*: No patients in either arm achieved a complete response. Partial response, 5 patients (19.2%; 95% CI, 8.3% to 30.1%) vs 5 patients (20.8%; 95% CI, 9.5% to 32.1%); progressive disease, 5 patients (19.2%) vs 5 patients (20.8%); stable disease, 16 patients (61.6%) vs 14 patients (58.4%), for GEM and COMB arms, respectively (significance levels not provided)
- *Survival*: Median survival duration: GEM arm, 37 weeks; COMB arm, 48 weeks (log-rank test, Breslow test and Tarone-Ware test not significant; $p = 0.65$)
- *Time to disease progression*: Median time to disease progression: GEM arm, 35 weeks; COMB arm, 34 weeks
- *Haematological toxicity*: Grades 3 and 4: leucopenia (3.7% and 0% vs 19.2% and 11.5%), thrombocytopenia (7.4% and 0% vs 7.7% and 0%), anaemia (7.4% and 0% vs 15.4% and 0%), for GEM and COMB arms, respectively. Febrile neutropenia occurred in 4 (15.4%) of patients in the COMB arm (1 treatment-related death due to septic shock)
- *Non-haematological toxicity*: In the GEM arm, 3.7% of patients experienced grade 3 nausea/vomiting. In the COMB arm, 23.1% and 11.5% of patients experienced grade 3 and grade 4 nausea/vomiting, respectively. There were no other grade 3 or 4 toxicities reported for the GEM arm. Grade 3 alopecia (3.8%), neurological events (3.8%) and grade 4 diarrhoea (3.8%) were observed in the COMB arm

Comments

Methodological comments

- Randomisation of eligible patients to GEM and COMB arms was performed by a statistical office not involved in the trial, using a computer-generated list of random numbers. Not placebo-controlled
- Baseline characteristics of age and gender were equally distributed in GEM and COMB arms. Histology, PS and staging were equally distributed across both arms
- Blinding was not specified
- ITT analysis was used for survival and toxicity analysis. Conducted log-rank, Breslow and Tarone-Ware tests for survival data. 'Assessable patients' only for evaluation of response (GEM arm, 26; COMB arm, 24)
- Sample size/power calculations were not specified
- Two COMB individuals were excluded from trial due to protocol violation (one had grade 3 hearing impairment before entry to trial, and one had brain metastases). In GEM arm, all patients received 2+ cycles, except two patients (one refused for no specific reason, and one died early due to disease progression). Two COMB patients and one GEM patient were not eligible for response evaluation

continued

General comments

- Inclusion and exclusion criteria were clearly described. Additional inclusion criteria included: prior malignancy excluded, with exception of *in situ* carcinoma of cervix or adequately treated basal cell carcinoma of skin; adequate bone marrow reserve (WBC, platelet, Hb, serum bilirubin and creatinine)
- Interventions were clearly described. Outcomes were clearly described
- All patients underwent baseline assessment of patient medical history, physical examination, PS, FBC, urinalysis, serum biochemistry profile, ECG, prothrombin time, activated partial thromboplastin time, chest roentgenography, whole-body bone scan, chest CT scan and brain CT scan (if clinically indicated)
- *Conflicts of interest:* None

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	2 COMB patients excluded from analysis because protocol violated (1 suffered grade 3 hearing impairment before entering trial, and other found to have brain metastases on second day of first cycle All GEM patients received 2 or more cycles, except 2 patients (1 refused without reason, and 1 died early of disease progression) 2 COMB patients and 1 GEM patient were not eligible for response evaluation

Reference and design	Intervention	Participants	Outcome measures
Sandler *et al.*, 2000[29] (International) Phase III, multicentre, randomised trial Funding: Eli Lilly Jadad score: 1/5	***GEM+CDDP arm*** GEM 1000 mg/m^2 (iv) on days 1, 8 and 15 of 28-day cycle, with CDDP 100 mg/m^2 on day 1 of 28-day cycle ***CDDP arm*** CDDP 100 mg/m^2 on day 1 of 28-day cycle ***Length of treatment*** Maximum of 6 cycles ***Other interventions used*** Pretreatment iv hydration	*n* = 522 GEM+CDDP arm: *n* = 260 CDDP arm: *n* = 262 71% of patients were male 69% of patients had stage IV NSCLC 84% of patients had good/excellent PS (80–100) 42% patients had adenocarcinoma 28% of patients had squamous cell carcinoma ***Characteristics of target population*** Histologically/cytologically confirmed stage IIIA or IIIB or IV NSCLC Median age: GEM+CDDP arm, 62 years (range, 36–88 years); CDDP arm, 63 years (range, 35–79 years) No prior chemotherapy/biological response-modifier therapy PS of 70–100 (Karnofsky scale) Measurable/assessable disease See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Measured from the time of randomisation ***Secondary end-points*** Response rate: SWOG response criteria, evaluated before each therapy cycle Toxicity: WHO criteria, evaluated at the beginning of each therapy cycle Time to disease progression: Measured from the time of randomisation QoL: HRQoL FACT-L, completed at baseline and the end of each cycle of therapy (only patients from the UK, USA and Canada) See *General comments* below ***Length of follow-up*** Not specified

Results

- *Survival*: Survival curves shown; GEM+CDDP curve above CDDP curve. Estimated median survival was significantly different (log-rank test, $p = 0.004$) between the two arms: GEM+CDDP arm, 9.1 months (95% CI, 8.3 to 10.6 months), vs CDDP, 7.6 months (95% CI, 6.5 to 8.2 months). One-year survival probability estimated at 39% for GEM+CDDP arm vs 28% for CDDP arm
- *Response*: GEM+CDDP arm (30.4%) had statistically higher ($p < 0.0001$) overall response rate than CDDP alone (11.1%). Complete response, 1.2% vs 0.4%; partial response, 29% vs 11%; progressive disease, 15% vs 33%; stable disease, 37% vs 42%; for GEM+CDDP and CDDP arms, respectively. Median number of treatment cycles: GEM+CDDP arm, 4 cycles; CDDP arm, 2 cycles
- *Time to disease progression*: GEM+CDDP curve above CDDP curve at all points. Estimated median time to progressive disease was significantly different between the two arms (log-rank test, $p = 0.0013$): GEM+CDDP arm, 5.6 months (95% CI, 4.6 to 6.1 months), vs CDDP arm, 3.7 months (95% CI, 3.3 to 4.2 months)
- *QoL*: Of the randomised patients, 72.4% (378) participated in HRQoL assessment (number who completed baseline and at least one other FACT-L questionnaire: GEM+CDDP arm, 161 patients; CDDP arm, 149 patients). Median number of HRQoL assessments completed: GEM+CDDP arm, 4; CDDP arm, 3. Baseline and median changes at last observation for each patient were not different between treatment arms. Patients in both arms noted decrease in physical, functional well-being and total FACT-L scores, but there was no difference in other subscales; the change was not statistically significant. Both arms noted a decrease in HRQoL, but there were no differences between arms
- *Haematological toxicity*: Grades 3 and 4: anaemia (21.8% and 3.2% vs 5.7% and 0.8%), thrombocytopenia (25% and 25% vs 2.8% and 0.8%), granulocytopenia (21.7% and 35.3% vs 3.3% and 1.2%), for GEM+CDDP and CDDP arms, respectively (significance levels not provided). Hospitalisation for febrile neutropenia: GEM+CDDP arm, 4.6%; CDDP arm, 1.3%
- *Non-haematological toxicity*: No significant difference between the two treatment arms. Grades 3 and 4: nausea (25% and 2% vs 20% and 1%), vomiting (11% and 12% vs 10% and 9%), creatinine (4.4% and 0.4% vs 1.6% and 0.4%), pulmonary dyspnoea (4% and 3% vs 3% and 2%), neuromotor events (11.5% and 0% vs 2.5% and 0%), for GEM+CDDP and CDDP arms, respectively. Asthenia and malaise (no grading scale reported): 58.0% and 4.2% in the GEM+CDDP arm, respectively, and 40.2% and 3.8% in the CDDP arm, respectively
- *Second-line chemotherapy received*: GEM+CDDP arm, 18%; CDDP arm, 27%. Receiving VNB as second-line chemotherapy: GEM+CDDP arm, 37%; CDDP arm, 40%. Receiving taxanes (DOC and PAX) as second-line chemotherapy: GEM+CDDP arm, 46%; CDDP arm, 31%. Receiving GEM as second-line chemotherapy: GEM+CDDP arm, 0%; CDDP arm, 14%

Comments

Methodological comments

- Trial was randomised, though no indication of method of randomisation. Not placebo-controlled
- Baseline characteristics (prognostic factors) of patients, in each arm, were comparable
- Blinding was not performed
- ITT analysis was used. Comparison of tumour response rates between arms was performed by Fisher's exact test. Kaplan–Meier analysis was performed using LIFETEST procedure in SAS. Cox proportional hazard analysis of data from time-to-event variables (e.g. survival time and time to progressive disease) were performed using PHREG procedure in SAS. Comparison of change in FACT-L scores between arms was performed using a non-parametric paired *t*-test and analysis of variance
- Sample size/statistical power was not specified. Concern exists that the power of the study to detect differences has not been addressed
- Attrition/dropouts from the study not specified. Concern exists that the number of patients withdrawing from the study was not reported
- Concern exists that the statistical difference for adverse events between arms was not reported
- Concern exists that non-English-speaking patients included in the study did not receive HRQoL questionnaire

continued

General comments

- Eligibility and inclusion criteria were clearly defined. Additional inclusion criteria included: prior radiation therapy was allowed as long as the irradiated area was not the only source of measurable disease and the therapy was completed ≤ 3 weeks before enrolment into study; adequate bone marrow reserve (WBC, platelets, Hb); adequate renal function; 4 weeks must have passed since major surgery; patients must have recovered from all toxicity; prior malignancy was allowed, provided patient had been disease-free for 5+ years; no prior CT or biological response-modifier therapy)
- Interventions were clearly defined. Outcome measures were clearly defined for response, toxicity and survival (response rate: complete response, the complete disappearance of all clinically detectable malignant disease and return of all abnormal tests to normal values for a period of at least 4 weeks; partial response, a decrease of at least 50% of the sum of the cross-sectional areas of all measured lesions in the absence of progression of any existing lesions for at least 4 weeks or any appearance of any new lesions within that time)
- FBC with differential and platelet count was performed on days of treatment. Serum creatinine was evaluated on first treatment day of each cycle. Before each course, past history, a physical examination and PS were documented, as were measurements of serum creatinine, hepatic transaminases and bilirubin, and urinalysis. If a chest X-ray was adequate to document bi-dimensionally measurable disease, this was repeated before each cycle of therapy. But, if chest or abdominal CT scans were required to document bi-dimensionally measurable disease, these were repeated at the beginning of every other cycle. HRQoL was evaluated using FACT-L at the end of each cycle of therapy
- *Conflicts of interest*: Study was part supported by Eli Lilly and Company, the Hoosier Oncology Group and the Walther Cancer Institute

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	0
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	

RTOG, Radiation Therapy Oncology Group; SD, standard deviation; ALT, alanine transaminase; AST, aspartate transaminase; WBC, white blood cell count; Hb, haemoglobin; AJCC, American Joint Commission for Cancer; TriCOMB, triple combination therapy; FBC, full blood count; COMB, combination therapy; LIFETEST, statistical procedure in Statistical Analysis Software; SAS, Statistical Analysis Software (SAS Institute, Cary, NC, USA); PHREG, statistical procedure in Statistical Analysis Software

Appendix 7

Summary of evidence of effectiveness of paclitaxel in lung cancer

Reference and design	Intervention	Participants	Outcome measures
Bonomi *et al.*, 2000[30] Phase III, multicentre, randomised trial (USA) Funding: NCI, National Institutes of Health, and Department of Health and Human Services Jadad score: 2/5	***Group 1*** VP-16 100 mg/m^2 (iv) on days 1–3, with CDDP 75 mg/m^2 (iv) on day 1. Each regimen was repeated every 21 days ***Group 2*** PAX 250 mg/m^2 (iv) on day 1, with CDDP 75 mg/m^2 (iv) on day 2. Each regimen was repeated every 21 days ***Group 3*** PAX 135 mg/m^2 (iv) on day 1, with CDDP 75 mg/m^2 (iv) on day 1. Each regimen was repeated every 21 days ***Length of treatment*** Not stated ***Other interventions used*** Filgrastim	*n* = 574 Group 1: *n* = 193 Group 2: *n* = 191 Group 3: *n* = 190 64% of patients were male 68% of patients had PS of 1 81% of patients had stage IV disease ***Characteristics of target population*** Histologically or cytologically confirmed NSCLC Bi-dimensionally measurable or assessable stage IIIB or IV disease Median age: Group 1, 61.7 years; Group 2, 60.8 years; Group 3, 62.7 years ECOG PS of 0 or 1 Allowed to have previous radiation, provided it had been completed ≥ 2 weeks before trial entry No previous chemotherapy allowed See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculation of survival time not specified ***Secondary end-points*** QoL: FACT-L (version 2) completed immediately before first course of chemotherapy, and again at 6, 12 and 26 weeks Response rate: ECOG criteria, evaluated before each therapy cycle Toxicity: ECOG criteria, evaluated before each therapy cycle ***Length of follow-up*** Median follow-up of 28.5 months See *General comments* below

Results

- *Survival*: Kaplan–Meier estimates were calculated after a median follow-up of 28.5 months and after 91.6% of the eligible patients had expired. Survival curves shown. Group 2 curve is above Group 1 curve at all points and comparable to Group 3 curve until approximately 17 months (Group 3 curve increases above Group 2 curve). Group 1 curve is below both PAX curves at all points. The combined PAX curve is above the VP-16 curve. There was marginally (non-significant) improved survival for each PAX arm compared with the VP-16 arm (log-rank comparisons: Group 1 vs Group 2, $p = 0.097$; Group 1 vs Group 3, $p = 0.9$). Survival between the two PAX regimens was not different (log-rank test, $p = 0.931$)
- *Survival duration and rates*: Median survival duration: Group 1, 7.6 months; Group 2, 10 months; and Group 3, 9.5 months. One-year survival rates: Group 1, 31.8%; Group 2, 40.3%; and Group 3, 37.4%. Median survival time for all patients receiving PAX: 9.9 months, with a 1-year survival rate of 38.9%. Median survival duration and 1-year survival rate for stage IIIB patients: Group 1, 7.9 months and 40.0%; Groups 2 and 3 combined, 13.1 months and 54.9%, respectively ($p = 0.152$). Median survival duration and 1-year survival rate for stage IV patients: Group 1, 7.6 months and 30.3%; Group 2 and 3 combined, 8.9 months and 34.5%, respectively ($p = 0.246$)
- The differences in survival for the individual treatment regimens (Group 1 vs Group 2 vs Group 3) in stage IIIB and IV patient subsets were not significant
- *Objective response rate (complete response plus partial response)*: Group 1, 12.4%; Group 2, 27.7%; and Group 3, 25.3%. Overall complete remission rate was 1.6%; the response rates for the PAX regimens were significantly higher than the VP-16+CDDP control regimen ($p < 0.001$ for Group 1 vs Group 2; $p = 0.002$ for Group 1 vs Group 3). Response rates for the two dose levels of PAX were not significantly different ($p = 0.264$)
- *QoL:* FACT-L completed by 94% of patients at baseline. The compliance rates among surviving patients at 6, 12 and 26 weeks were 72%, 60%, and 50%, respectively. The change in TOI and FACT-L scores, over 6 months, for each regimen (9.3 and 11.8 vs 10.2 and 11.4 vs 8.3 and 9.9) for Group 1, Group 2 and Group 3, respectively. Table given for short-term (baseline and 6 weeks) QoL responses. A higher percentage of PAX patients were classified as having improved QoL, but difference was not significant in the analysis, which excluded patients who failed to complete the 6-week FACT-L ($p = 0.46$). However, if patients with missing data were included, then there would be a significantly higher rate of improved QoL for PAX patients ($p = 0.012$)

continued

***Results* contd**

- *Haematological toxicity*: Grade 4 leucopenia (16.0% vs 27.0% vs 14.0%), grade 4 granulocytopenia (55.0% vs 65.0% vs 74.0%), ≥ grade 3 infection (8.5% vs 9.0% vs 7.4%), grade 5 infection (1.0% vs 3.0% vs 1.5%), grade 4 thrombocytopenia (5.0% vs 5.0% vs 0.5%), ≥ grade 3 anaemia (28.0% vs 19.0% vs 21.0%), for Group 1, Group 2 and Group 3, respectively
- *Non-haematological toxicity*: Grade 3 neurological events (21.0% vs 40.0% vs 23.0%); grade 4 nausea/vomiting (6.0% vs 9.0% vs 10.0%); ≥ grade 3 myalgias (0.0% vs 7.0% vs 1.0%); grade 5 cardiac events (possibly treatment related) (0.5% vs 0.5% vs 2.0%), with significantly more serious and possibly treatment-related cardiac events in Group 2 ($p = 0.026$) but not in Group 3 ($p = 0.13$); grade 5 all types (1.5% vs 5.0% vs 4.0%), for Group 1, Group 2 and Group 3, respectively. Grade 5 cardiac events included sudden death in 3 patients, myocardial infarction in 2 patients and hypotension associated with acute pericarditis in 1 patient. There was a history of cardiovascular disease in 4 of the patients who died as a result of cardiovascular disease: known coronary artery disease (2), hypertension (1) and cardiac arrhythmia (1). One patient who died suddenly had no known risk factors. Three additional deaths seemed treatment related: respiratory insufficiency (2; 1 in each PAX arm), renal insufficiency (1; in high-dose PAX arm)

Comments

Methodological comments

- Patients were randomised to one of the three regimens. Stratification occurred based on PS (0 vs 1), weight loss during previous 6 months (< 5% vs ≥ 5%), stage IIIB vs IV disease, and bi-dimensionally measurable disease vs assessable disease. The control regimen was defined as Group 1 (VP-16+CDDP)
- Baseline characteristics between the three treatment groups were comparable
- Blinding was not specified (although the article states that interim analyses based on primary efficacy comparison were performed)
- Fisher's exact test was used to compare response rates, and the Kruskal-Wallis test was used to compare degrees of toxicity. Survival estimates were calculated by the Kaplan–Meier method, and the log-rank test was used for survival comparisons. A joint mixed effects and survival model was used to estimate change in QoL scores as a continuous variable over 6 months and to compare the three regimens. Eligible patients were analysed
- The accrual goal of the study (585 patients) was estimated to have more than 90% power to detect a 50% increase in median survival time from 6 months in patients on the VP-16+CDDP regimen to 9 months in patients on either of the PAX+CDDP regimens, at an overall, experiment-wise significance level of 5% to be monitored for up to four times using an O'Brien-Fleming-type group sequential method
- Of the 599 patients first randomised in the study, 11 were withdrawn, for the following reasons: withdrawal of consent (5), death before start of treatment (3), discovery of brain metastases before starting protocol treatment (2) and infection (1). Also, 14 patients were classified as ineligible: PS ≥ 2 (5), required tests performed > 2 weeks before protocol entry (2), multiple primary cancers (2), inadequate organ function (1), less than stage IIB disease (1), radiation given to only measurable site (1), small-cell histology (1) and radiation given concurrently with start of protocol therapy (1)

General comments

- Inclusion and exclusion criteria were clearly defined. Additional inclusion criteria included: no previous history of malignant disease, with the exception of skin cancer or carcinoma *in situ* of the cervix; adequate organ function; no active infections or brain metastases; patients with only measurable lesion within a previously radiated field; patients could not have uncontrolled diabetes mellitus, uncontrolled hypertension, unstable angina, congestive heart failure, myocardial infarction within previous year or evidence of pre-existing peripheral neuropathy)
- Interventions were clearly defined. Outcome measures relating to response, toxicity and QoL were clearly defined (QoL assessed using version 2 of the FACT-L instrument, which consists of 35 questions, each of which is scored using a 5-point scale; the questions address six areas: physical well-being, functional well-being, lung cancer symptoms and concerns, social well-being, emotional well-being, and relationship with the doctor; 21 questions on physical and functional well-being and lung cancer symptoms were combined and designated as the TOI, which is considered to be the best summary indicator of the physical component of QoL). Calculation of survival time was unclear/not specified
- Before each treatment cycle, history and physical examination, FBC, serum chemistries and tumour measurements were performed; QoL questionnaire was filled in immediately before receiving first chemotherapy course, and again 6, 12 and 26 weeks later
- *Conflicts of interest*: Supported by Public Health Service Grants from the NCI, National Institutes of Health, and Department of Health and Human Services

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	11/599 patients (599 were first randomised in study) were withdrawn, for following reasons: withdrawal of consent (5), death before start of treatment (3), discovery of brain metastases before starting protocol treatment (2) and infection (1) 14 patients were classified as ineligible: PS ≥ 2 (5), required tests performed > 2 weeks before protocol entry (2), multiple primary cancers (2), inadequate organ function (1), less than stage IIB disease (1), radiation given to only measurable site (1), small-cell histology (1) and radiation given concurrently with start of protocol therapy (1)

Reference and design	Intervention	Participants	Outcome measures
Chang *et al*., 1993[31] (USA) Phase II, randomised trial Funding: NCI Jadad score: 2/5	***PAX arm*** PAX 250 mg/m^2 (iv), given in doses of 250 mg/m^2, repeated every 3 weeks ***MER arm*** MER 1000 mg/m^2 (iv), repeated every 3 weeks ***PIR arm*** PIR 150 mg/m^2 (iv), repeated every 3 weeks Doses were modified or withdrawn on haematological, hepatic, cardiac and renal toxicity ***Length of treatment*** Not specified ***Other interventions used*** Sodium chloride, dextrose, dexamethasone, cimetidine, diphenhydramine, glucose	*n* = 113 PAX arm: *n* = 25 MER arm: *n* = 35 PIR arm: *n* = 44 68% of patients were male 63% of patients had PS of 1 54% of patients had adenocarcinoma ***Characteristics of target population*** Histologically confirmed diagnosis of NSCLC, with clinical evidence of stage IV measurable disease Median age: PAX arm, 61 years (range, 38–82 years); MER arm, 62 years (range, 42–80 years); PIR arm, 61 years (range, 31–85 years) PS of 0 or 1 (ECOG criteria) No prior chemotherapy No brain metastases See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: ECOG response criteria, evaluated every 3 weeks ***Secondary end-points*** Toxicity: ECOG and NCI criteria, evaluated every 3 weeks Time to treatment failure: Measured from entry into study to the time of relapse, progression or death, or to last known date that patient was alive Survival: Definition of calculation of survival time not provided See *General comments* below ***Length of follow-up*** Median follow-up was > 17 months

Results

- *Response*: Complete response: 1/44 patients (2.3%) recorded in PIR arm. Partial response: 5/24 patients (20.8%; 95 % CI, 7% to 42%) in PAX arm and 2/35 patients (5.7%) in MER arm. Response duration of patients with partial response: PAX patients, 3.7 months, 5 months, 6.4 months, > 6.5 months and > 15.4 months; MER patients, > 9 months and 3.6 months; PIR patient, > 6 months. Stable disease: PAX arm, 5/24 patients (20.8%); MER arm, 4/35 patients (11.4%); PIR arm, 6/44 patients (13.6%)
- *Time to treatment failure*: Median time to treatment failure: PAX arm, 8.9 weeks; MER arm, 6.9 weeks; PIR arm, 9.1 weeks
- *Survival*: Survival curves shown: roughly equivalent until 30 weeks, with PAX arm above other arms after this point. PAX arm, 19/24 eligible patients (79.2% died); MER and PIR arms, 67/79 patients (85.7%) died. Median survival time: PAX arm, 24.1 weeks; MER arm, 19.9 weeks; and PIR arm, 29.3 weeks (NS). One-year survival rates: PAX arm, 41.7% ± 10% (SE); MER arm, 21.6% ± 7%; PIR arm, 22.6% ± 7%. All 5 responders in PAX arm and 2/5 patients with stable disease on PAX survived longer than 1 year. The survival difference at 1 year was not statistically significant because of small sample size
- *Time to treatment failure curve*: PAX curve above MER and PIR curves > 15 weeks after follow-up
- *Haematological toxicity*: Grades 3 and 4: leucopenia (83% vs 6% vs 23%), thrombocytopenia (4% vs 0% vs 0%), anaemia (21% vs 14% vs 4.5%), infection (12.5% vs 3% vs 5%), for PAX, MER and PIR, respectively
- *Non-haematological toxicity*: Grades 3 and 4: nausea/vomiting (4% vs 17% vs 0%), diarrhoea (8% vs 0% vs 0%), stomatitis (4% vs 0% vs 0%), hepatic (0% vs 6% vs 9%), cardiac (17% vs 6% vs 0%), neurosensory (8% vs 0% vs 0%), neuromotor (12.5% vs 14% vs 0%), neuropsychological (4% vs 8.5% vs 0%), neuroclinical (4% vs 11% vs 4.5%), pulmonary (0% vs 6% vs 0%), genitourinary (0% vs 6% vs 0%), phlebitis (0% vs 3% vs 0%), for PAX, MER and PIR, respectively. Grade 5 infection (4% vs 3% vs 0%), grade 5 cardiac (0% vs 3% vs 2%), grade 5 pulmonary (0% vs 3% vs 2%), grade phlebitis (0% vs 3% vs 0%), for PAX, MER and PIR, respectively
- *Toxic deaths*: PAX, 1 (4%), due to sepsis; MER, 4 (11.4%), due to thrombosis, sepsis, cardiac collapse and pulmonary emboli; PIR, 2 (5%), due to myocardial infarction and respiratory failure; in all arms, deaths were directly related to treatment. Atrial fibrillation occurred in 2 PAX patients. One PAX patient had myocardial infarction, and 1 PAX patient had mild congestive heart failure, with orthopnoea and abnormal LVEF. In PAX arm, 1 patient developed depression, 2 developed grade 3 sensory neuropathy, and 3 developed grade 3 motor neuropathy. These were thought not to be directly due to PAX

Comments

Methodological comments

- Patients were randomised equally among the three regimens until the PAX accrual reached 25 (due to a limited supply of PAX at time of study); thereafter, patients were randomised equally between MER and PIR arms
- There was no difference in the distribution of prognostic factors between the three treatment arms
- Blinding was not specified
- Not ITT. Analysis of survival and time to treatment failure was based on the Kaplan–Meier method
- Statistical design was planned to have a large enough patient number in each treatment arm to guarantee reasonable accuracy in the estimated response rate. The study entered 40 evaluable patients in MER and PIR arms, thus the 95% CI of a given response rate would be no wider than 32%
- Five patients in the MER arm and 1 in the PIR arm were excluded because of: problems with drug supplies (1), insurance company's refusal to pay for treatment (2), brain metastases (2) and death before treatment started (1). Nine others were deemed ineligible: 1 patient on PAX (lack of metastatic disease), 2 patients on PIR (second malignancy [1] and no pretreatment measurement within 2 weeks of registration [1]), and 6 patients on MER (without measurable disease [1], PS of 2 [2], no pretreatment measurement within 2 weeks of registration [2]). One additional patient on MER was excluded from analysis because of inadequate data

continued

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate haematological, renal and hepatic functions. Exclusion criteria: WBC < 4000/µl, platelet count < 100,000/µl, bilirubin level > 1.5 mg/dl, random AST level > twice normal, creatinine level > 1.5 mg/dl, random blood sugar level > 200 mg/dl, abnormal ejection fraction of the left ventricle, history of myocardial infarction within 12 months of study entry, uncontrolled hypertension, heart failure, arrhythmias or prior history of malignancy other than skin cancer that had been cured)
- Interventions were clearly defined. Outcome measures were clearly defined
- Patients were evaluated with history and physical examination, complete blood counts, measurement of tumour size, and toxicity. Weekly assessment included blood counts (to monitor myelosuppression), creatinine levels (for MER patients) and serum transferase levels (for PIR patients). The case history, study forms and radiographic evaluations of the responders treated with PAX were independently reviewed by representatives from the NCI, ECOG and Bristol-Myers Squibb Company
- *Conflicts of interest*: Supported by Public Health Service Grants from the NCI, National Institutes of Health, and Department of Health and Human Services

Quality assessment for RCTs (Jadad score[19])

Question	**Score**
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	5 patients in MER arm and 1 patient in PIR arm were excluded because of: problems with drug supplies (1), insurance company's refusal to pay for treatment (2), brain metastases (2) and death before treatment started (1) 9 others were deemed ineligible: 1 patient on PAX (lack of metastatic disease), 2 patients on PIR (second malignancy [1] and no pretreatment measurement within 2 weeks of registration [1]), and 6 patients on MER (without measurable disease [1], PS of 2 [2], no pretreatment measurement within 2 weeks of registration [2]) 1 additional patient on MER was excluded from analysis because of inadequate data

Reference and design	Intervention	Participants	Outcome measures
Giaccone *et al.*, 1998[35] (The Netherlands) Phase III, randomised trial Funding: Bristol-Meyers Squibb International Jadad score: 3/5	***CDDP+VM-26 arm*** CDDP 80 mg/m^2 (iv) on day 1, with VM-26 100 mg/m^2 (iv) on days 1, 2 and 5. Cycle repeated every 3 weeks ***PAX+CDDP arm*** CDDP 80 mg/m^2 (iv) on day 1, with PAX 175 mg/m^2 (iv) on day 1. Cycle repeated every 3 weeks ***Length of treatment*** Maximum of 6 cycles, in absence of disease progression and unacceptable toxicity ***Other interventions used*** Saline, 5% dextrose, fluids, 5% glucose Prophylactic medication prescribed: oral dexamethasone, diphenhydramine, cimetidine Antiemetics: ondansetron, dexamethasone	*n* = 332 patients PAX+CDDP arm: *n* = 166 CDDP+VM-26 arm: *n* = 166 71% of patients were male 71% of patients had < 5% weight loss 89% of patients had PS of 0 or 1 51% of patients had adenocarcinoma 61% of patients had stage IV disease ***Characteristics of target population*** Cytological or histological diagnosis of locally advanced or metastatic NSCLC Aged between 18 and 75 Median age: CDDP+VM-26 arm, 58.5 years (range, 28–75 years); PAX+CDDP arm, 58.5 years (range, 34–74 years) No prior chemotherapy PS < 3 (WHO criteria) Assessable or measurable disease required Prior radiotherapy was allowed if at least 4 weeks before study entry and if not all marker lesions were irradiated See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculated from the date of randomisation ***Secondary end-points*** Progression-free survival: Calculated from the date of randomisation Response rate: WHO criteria, evaluated every 2 cycles of chemotherapy QoL: EORTC QLQ-C30-LC13, administered at baseline and every 6 weeks thereafter, for four times (assessment performed at selected institutions) Toxicity: NCIC-CTC, evaluated before each therapy cycle See *General comments* below ***Length of follow-up*** Median follow-up duration of 21 months

Results

- At time of analysis, 244 patients (CDDP+VM-26 arm, 124 patients; PAX+CDDP arm, 120 patients) had died after median follow-up duration of 21 months (randomisation between July 1993 and February 1996)
- *Survival*: Survival curves shown; two curves highly comparable. Median survival time: CDDP+VM-26 arm, 9.9 months; PAX+CDDP arm, 9.7 months ($p = 0.971$). One-year survival rate: CDDP+VM-26 arm, 41% (95% CI, 33% to 49%); PAX+CDDP arm, 43% (95% CI, 25% to 51%). Two-year survival rate: CDDP+VM-26 arm, 18% (95% CI, 10% to 26%); PAX+CDDP arm, 19% (95% CI, 12% to 26%). Median progression-free survival: CDDP+VM-26 arm, 4.9 months; PAX+CDDP arm, 5.4 months ($p = 0.973$). One-year progression-free survival: CDDP+VM-26 arm, 18% (95% CI, 12% to 24%); PAX+CDDP arm, 17% (95% CI, 11% to 23%). Two-year progression-free survival: CDDP+VM-26 arm, 8% (95% CI, 3% to 13%); PAX+CDDP arm, 6% (95% CI, 1% to 11%). Multivariate Cox model (following variables retained at 5% level of significance): weight loss ($p = 0.013$), PS ($p < 0.001$), extent of disease ($p = 0.007$), thrombocytosis ($p < 0.001$) and alkaline phosphatase ($p = 0.001$); taking these variables into account, the treatment effect was not significant ($p = 0.239$)
- *Response*: Complete response, 1 (1%) vs 2 patients (1%); partial response, 44 (27%) vs 61 patients (39%); progressive disease, 29 (18%) vs 20 patients (13%); stable disease, 67 (41%) vs 58 patients (37%), for CDDP+VP-26 and PAX+CDDP, respectively. Overall response rate was significantly higher ($p = 0.018$) for the PAX+CDDP arm (41%; 95% CI, 33% to 48%) than for the CDDP+VP-26 arm (28%; 95% CI, 21% to 35%). Median duration of response was significantly higher ($p = 0.362$) for the PAX+CDDP arm (8.3 months) than for the CDDP+VM-26 arm (9.5 months). Analysis of clinical response by logistic regression (following variables retained in model at 5% level of significance): age (higher response in older patients; $p = 0.029$), PS (response rate, 37% vs 12% for PS of 0–1 vs 2, respectively; $p = 0.008$) and treatment (in favour of PAX+CDDP; $p = 0.012$)
- *QoL*: In total, 104 patients completed the baseline questionnaire. Compliance to the administration of the QoL questionnaire decreased over time. For most of the functioning scales, there was a significant benefit in favour of the PAX+CDDP arm at the 6-week assessment ($n = 94$) on the symptom scales; fatigue ($p = 0.006$) and appetite loss ($p < 0.001$) were significantly lower at 6 weeks in the PAX+CDDP arm. However, this benefit was no longer significant at the 12-week assessment ($n = 74$). In contrast, at 12 weeks, peripheral neuropathy was reported in 26 patients (76%) in PAX+CDDP arm compared with 5 patients (18%) in CDDP+VM-26 arm. A logistic regression model indicated that patients with a poorer PS tended to drop out of the study earlier
- *Haematological toxicity*: Grades 3 and 4: leucopenia (34% and 32% vs 16% and 3%; $p < 0.001$), neutropenia (16% and 67% vs 27% and 28%; $p < 0.001$), thrombocytopenia (18% and 18% vs 1% and 1%; $p < 0.001$), anaemia (21% and 3% vs 9% and 1%; $p = 0.002$), for CDDP+VP-26 and PAX+CDDP arms, respectively. Febrile neutropenia (27% vs 3%; $p < 0.001$)), for CDDP+VP-26 and PAX+CDDP arms, respectively
- *Non-haematological toxicity*: Grades 3 and 4: infection (10% and 4% vs 3% and 0%; $p = 0.02$), for CDDP+VP-26 and PAX+CDDP arms, respectively. Grade 3: stomatitis (1% vs 1%), myalgia/arthralgia (1% vs 3%), peripheral neurotoxicity (1% vs 9%), for CDDP+VP-26 and PAX+CDDP arms, respectively. Cardiac toxicity and other non-haematological toxicities were similar in both arms, apart from alopecia, which was present in the majority of patients, and vomiting, which was severe in 12% of patients overall; the other side-effects were infrequent. A table or breakdown by treatment arm was not provided

continued

Comments

Methodological comments

- Randomisation was performed centrally by the EORTC Data Centre; patients were stratified by institution, PS (0 vs 1 or 2) and extent of disease (locally advanced vs metastatic)
- Patient characteristics were well balanced between arms. Baseline characteristics of those completing questionnaire were well balanced and were similar to the entire patient population
- An independent team of experts, blinded with respect to treatment, reviewed responses
- All analyses were performed on all eligible patients, using ITT analysis. The Kruskal-Wallis exact test was used to compare characteristics in both arms. Survival curves were estimated using the Kaplan–Meier method, and differences were assessed by the log-rank test at the 0.05 level of significance. To adjust for confounding variables, the Cox proportional hazards model with backward variable selection procedure was used
- Assuming a median survival duration of 7 months in CDDP+VM-26 arm, a total of 248 deaths was necessary to detect an absolute increase of 3 months in the median survival time (i.e. 10 months in PAX+CDDP arm) with a two-sided type I error of 0.05 and a power of 80%. A total of 280 patients was necessary, assuming a recruitment period of 3 years and 1 year of follow-up evaluation. Multivariate model was based on: age, sex, PS, weight loss in 3 months before randomisation, histology, metastases and pretreatment values of chemistries. Regression analysis was also applied to investigate influence of treatment and prognostic factors on the clinical response. Two-sided Wilcoxon rank-sum test with stratification for sex was used to compare QoL scores (sex used for stratification because differences at baseline were observed)
- Fifteen patients were found to be ineligible after randomisation (4 in CDDP+VM-26 arm and 11 in PAX+CDDP arm) due to previous malignancy (8), brain metastases (2), all assessable lesions irradiated (4) and other malignancy (1). Of the 317 eligible patients, 5 patients never started treatment because of refusal (4) or rapid progression (1)

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: no second malignancy, except radically treated *in situ* carcinoma of the cervix or basal call carcinoma of the skin; adequate haematological reserve (granulocytes $\geq 2.0 \times 10^9$/l, platelets $\geq 100 \times 10^9$/l); normal hepatic function (bilirubin level ≤ 1.25 times upper normal value) and renal function (serum creatinine concentration ≤ 120 µmol/l and/or creatinine clearance ≥ 60 ml/minute); no signs of cardiac failure or rhythm disturbances that necessitated medication; no signs or symptoms of brain metastases
- Interventions were clearly defined. Outcome measures were clearly defined
- At study entry, full medical history and physical examination, FBC with differential chemistries, creatinine clearance, ECG and chest X-ray were performed. Before start of every cycle, study entry investigations were repeated. Response evaluation included CT scans, and ultrasounds were performed to document disease extent optimally and to evaluate response to treatment. CT scan of skull was not routinely required but was performed in the presence of neurological symptoms. Weekly blood cell counts were performed
- Concern exists for lack of information related to non-haematological toxicity
- *Conflicts of interest*: Supported by Bristol-Meyers Squibb International

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	15 (4.5%) of patients were found to be ineligible after randomisation (4 in CDDP+VM-26 arm and 11 in PAX+CDDP arm) due to previous malignancy (8), brain metastases (2), all assessable lesions irradiated (4) and other malignancy (1) Of the 317 eligible patients, 5 (1.6%) of patients never started treatment because of refusal (4) or rapid progression (1)

Reference and design	Intervention	Participants	Outcome measures
Postmus *et al.*, 1996[33] (The Netherlands–Belgium) Phase II, multicentre, randomised trial (interim analysis) Funding: No funding information provided Jadad score: 2/5	***CDDP+VM-26 arm*** CDDP 80 mg/m^2 (iv) on day 1, with VM-26 100 mg/m^2 (iv) on days 1, 3, and 5. Course repeated every 3 weeks. CDDP given before VM-26 ***PAX+CDDP arm*** CDDP 80 mg/m^2 (iv) on day 1, with PAX 175 mg/m^2 (iv) on day 1. Course repeated every 3 weeks. PAX given before CDDP ***Length of treatment*** Maximum of 6 courses ***Other interventions used*** Saline, dextrose, fluids, prophylactic antiemetics, oral dexamethasone, diphenhydramine, cimetidine	*n* = 73 CDDP+VM-26 arm: *n* = 38 PAX+CDDP arm: *n* = 35 71% of patients were male 86% of patients had PS of 0 or 1 53% of patients had adenocarcinoma 64% of patients had stage IV disease ***Characteristics of target population*** Histologically or cytologically proven NSCLC Locally advanced stage IIIB or IV, recurrent or widespread or not amenable to radical surgery Aged between 18 and 76 years Median age: CDDP+VM-26 arm, 58 years (range, 28–75 years); PAX+CDDP arm, 59 years (range, 34–71 years) No previous chemotherapy WHO PS ≤ 2 See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: Investigations included chest X-ray, chest CT, CT/ultrasound of liver and other investigations (not specified) deemed necessary to document disease extent and follow response. Evaluated after 2, 4 and 6 therapy cycles ***Secondary end-points*** Toxicity: Evaluated after each therapy cycle (criteria not provided) Time to progression: Criteria not provided Time to death: Criteria not provided ***Length of follow-up*** Not specified See *General comments* below

Results

- At time of this interim analysis (date not provided), 10 patients were still alive (randomisation between July 1993 and October 1994)
- *Response*: Complete response, 0% vs 3%; partial response, 26% vs 37%; progressive disease, 24% vs 14%; no change, 39% vs 37%; for CDDP+VM-26 and PAX+CDDP arm, respectively (significance levels not provided). Not assessable: 2 patients in CDDP+VM-26 arm and 1 patient in PAX+CDDP arm
- *Early death*: Toxicity (2 patients vs 1 patient), other causes (0 patients vs 1 patient), for CDDP+VM-26 and PAX+CDDP arms, respectively
- *Haematological toxicity*: Main toxic events were grade 3-4 leucopenia (71% vs 23%), grade 3–4 neutropenia (82% vs 58%), grade 3–4 thrombocytopenia (37% vs 3%), grade 3–4 anaemia (19% vs 9%), febrile neutropenia (34% vs 3%), for CDDP+VM-26 and PAX+CDDP arms, respectively. No significance levels were provided
- *Non-haematological toxicity*: Grade 3 nausea (18% vs 11%), grade 3–4 vomiting (18% vs 15%), grade 3–4 diarrhoea (6% vs 0%), grade 2 stomatitis/mucositis (11% vs 0%), grade 2–3 infection (19% vs 6%), hypersensitivity reaction (3% vs 3%), grade 1–3 myalgia (10% vs 48%), grade 2–3 renal events (14% vs 9%), grade 3–4 cardiac events (8% vs 6%), grade 1–3 neurological events (60% vs 77%), for CDDP+VM-26 and PAX+CDDP arms, respectively. No significance levels were provided

Comments

Methodological comments

- Patients were randomised to two groups. During randomisation, patients were stratified according to the institution, PS (0 vs 1 or 2) and extent of disease (stage III vs IV)
- Prognostic factors between treatment arms were comparable. However, 60% of CDDP+VM-26 arm were adenocarcinoma patients, while only 45% of PAX+CDDP arm were adenocarcinoma patients. However, no tests of significance have been carried out
- Blinding was not specified
- Sample size/statistical power of study was unclear. Each treatment group had to include 26 evaluable patients if one assumed the regimen to be active, whereas in reality the response rate is uninteresting (< 10%) at $p = 0.10$, when a probability of 0.05 would mean rejecting a regimen with a high response rate (> 30%) from further study. If the number of responses from the first stage of the study was less than four, the trial had to be terminated and the regimen rejected. At least six responses were needed in each treatment arm to consider it sufficiently effective to warrant further study. If both regimens were to be considered to be sufficiently effective, the study was to be continued as a Phase III study
- *Attrition/dropout rate*: Seven patients were considered to be ineligible (2 in CDDP+VM-26 arm and 5 in PAX+CDDP arm); the reasons for withdrawal were previous malignancy (3), SCLC (1), all lesions included in prior radiotherapy field (2) and brain metastases (1). One patient dropped out after randomisation due to treatment refusal. There were 16 (42%) patients in CDDP+VM-26 arm and 17 (49%) in PAX+CDDP arm who discontinued due to disease progression. Three patients refused further treatment; toxicity was the reason for 9 (24%) patients in CDDP+VM-26 arm and 7 (24%) in PAX+CDDP arm to stop treatment

continued

General comments

- *Note*: Reported trial is an interim analysis of a Phase II trial, which would be developed as a Phase III trial if the results suggested continuation
- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: history of previous radiotherapy allowed if it had been discontinued 4 weeks prior to entry and target lesion was outside the radiation field; no brain metastases; no history of previous malignancy except for *in situ* carcinoma of the cervix or basal or squamous cell carcinoma of the skin
- Interventions were clearly defined. Outcome measures were not clearly defined. Concern exists that time to death and time to progression were suggested as end-points of the study, although no results of these end-points are reported
- Enrolment assessment included full medical history, physical examination, and documented height, weight and weight loss during the previous 6 months. This assessment also included WHO PS, biopsy or cytology confirming NSCLC, complete blood cell count including differential and platelets, chemistries, creatinine clearance, ECG, chest X-ray, CT scan of chest if useful to better document disease extent and follow response. Before every treatment course, physical examination, full blood cell counts with differentials and platelets, creatinine clearance, chest X-ray and ECG were performed
- *Conflicts of interest*: Not specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	7/80 (9%) patients were considered to be ineligible (2/40 [5%] in CDDP+VM-26 arm and 5/40 [12.5%] in PAX+CDDP arm; the reasons for withdrawal were previous malignancy (3), SCLC (1), all lesions included in prior radiotherapy field (2) and brain metastases (1) One patient dropped out after randomisation due to treatment refusal 16 (42%) patients in CDDP+VM-26 arm and 17 (49%) in PAX+CDDP arm discontinued due to disease progression Three patients refused further treatment Toxicity was the reason for 9 (24%) patients in CDDP+VM-26 arm and 7 (24%) in PAX+CDDP to stop treatment

Reference and design	Intervention	Participants	Outcome measures
Ranson *et al.*, 2000[32] Phase III, multicentre, randomised trial Funding: No information provided Jadad score: 3/3	***PAX+BSC arm*** PAX 200 mg/m^2 (iv) administration with cycles repeated every 21 days, plus best supportive care ***BSC arm*** BSC included palliative radiotherapy for bronchial obstruction, haemoptysis, superior vena caval obstruction and brain metastases. Cortico-steroids, antibiotics, analgesics, antiemetics, transfusions and other symptomatic therapy given as required ***Length of treatment*** Until disease progression or unacceptable toxicity ***Other interventions used*** PAX+BSC patients given 20 mg dexamethasone, 300 mg cimetidine or 50 mg ranitidine, and 50 mg diphenhydramine or 10 mg chlorpheniramine pretreatment	*n* = 157 PAX+BSC arm: *n* = 79 BSC arm: *n* = 78 75% of patients were male 55% of patients had stage IV disease 83% of patients had PS 0 or 1 47% of patients had squamous cell carcinoma 30% of patients had adenocarcinoma ***Characteristics of target population*** Histologically proven NSCLC Bi- or uni-dimensionally measurable stage IIIB or IV NSCLC Aged ≥ 18 years Median age: PAX+BSC arm, 65 years (range, 37–78 years); BSC arm, 64 years (range, 23–82 years) Life expectancy of ≥ 12 weeks PS of 0–2 (ECOG criteria) Previous radiotherapy was allowed but had to be outside assessable lesions, and used to treat < 30% of marrow-bearing bones and > 2 months prior to enrolment (this period was reduced to 4 weeks after 1 year into recruitment) Patients had to be English-speaking to participate in QoL questionnaire See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculated from the date of randomisation until the date of death ***Secondary end-points*** QoL: RSCL, completed at baseline, every 3 weeks during the study, upon leaving the study and thereafter at 6-week intervals. Continued until < 30% of the randomised patients were available for assessment Time to disease progression: Two analyses performed. Protocol definition of progression (not provided) included the need for subsequent radiotherapy. Also performed with patients who received radiotherapy prior to tumour progression being censored at the start of radiotherapy Toxicity: WHO criteria, evaluated every 3 weeks Response: Calculation definition not provided ***Length of follow-up*** After treatment discontinuation or reaching 8 months in the study, patients were assessed every 6 weeks until death (data set closure on 31 March 1998)

Results

- At the closure of the study (31 March 1998; randomisation between February 1995 and October 1997), one patient had been lost to follow-up, and 15 patients in the PAX+BSC arm and 7 in the BSC arm remained alive
- *Survival*: Survival curves shown; PAX+BSC curve above BSC curve. Median survival in PAX+BSC arm was 6.8 months (95% CI, 5.7 to 10.2 months) vs 4.8 months (95% CI, 3.7 to 6.8 months) in the BSC arm (log-rank test, $p = 0.037$). One-year survival 95% CIs were 20% to 41% for the PAX+BSC arm and 18% to 39% for the BSC arm. Stratified Cox regression model: PAX+BSC was significantly ($p = 0.048$) associated with increased survival, with a hazard ratio of 0.68 (95% CI, 0.49 to 1.0)
- *Time to progression*: At the point of analysis, 74/79 PAX+BSC patients and 77/78 BSC patients had progressive disease. Median time to disease progression was significantly longer ($p = 0.0001$) in the PAX+BSC arm at 3.9 months (95% CI, 3.3 to 4.5 months) compared with 0.5 months (95% CI, 0.4 to 0.7 months) in the BSC arm. Of the patients who received radiotherapy (64% in the BSC arm and 9% in the PAX+BSC arm), median time to disease progression was 4 months (95% CI, 3.4 to 5.3 months) in the PAX+BSC arm and 1.2 months (95% CI, 1.0 to 1.7 months). This result was statistically significant ($p = 0.0001$)
- *Response*: In the PAX+BSC arm, 3/79 patients were unavailable for assessment of response. There were 12/76 responses (1 complete response and 11 partial responses) for an overall response rate of 16% (95% CI, 8% to 26%). Response rates were 18% and 13% for stage IIIB and IV patients, respectively. None of the 13 patients with an ECOG response status of > 1 was a responder, nor were the 7 patients with large-cell carcinoma. The response rate for ECOG status 0–1 was 19%. Response was higher in older patients (21% vs 11% in those < 65 years) and higher in those who had not received radiotherapy (25% vs 14%). None of these percentages were statistically significant
- *QoL*: QoL analysis continued until week 33 (protocol dictated that 30% of patients remained available for QoL assessment). Compliance with the questionnaire was 95% and 96% at baseline and 64% and 60% at week 33, for the PAX+BSC and BSC arms, respectively. Random effects model analysis (on each of the five RSCL subscores) was performed on 'dropouts' (no QoL data after week 15) and 'completers' (QoL data beyond week 15). Slopes were smaller for the dropouts than the completers for each of the subscores, indicating that completers had more favourable QoL scores (estimated slope scores provided, coefficients not provided). The analysis indicated that, compared with BSC, improvement in the functional activity subscore of RSCL favoured the PAX+BSC arm for dropouts ($p = 0.043$). For all other subscores, there were no statistically significant differences between the two arms
- *Haematological toxicity*: Grade 3–4 adverse events: neutropenia (34% vs 0%), leucopenia (9% vs 0%), anaemia (3% vs 2%), for PAX+BSC and BSC arms, respectively
- *Non-haematological toxicity*: Grade 3–4 adverse events: alopecia (76% vs 0%), arthralgia/myalgia (22% vs 4%), asthenia (14% vs 9%), nausea/vomiting (5% vs 1%), infection (10% vs 3%), for PAX+BSC and BSC arms, respectively
- Patients required hospitalisation more frequently in the PAX+BSC arm (58%) than the BSC arm (41%); however, patients remained in the study a median 1.5 times longer in the PAX+BSC arm. Median number of PAX+BSC courses was 5 (range, 1–10), and median dose was 923 mg/m^2. Haematological and non-haematological toxic effects led to a dose reduction in 25/78 PAX+BSC patients (1 patient withdrew prior to chemotherapy) or 11% of cycles (39/357)

continued

Comments

Methodological comments

- Randomisation was performed centrally by Bristol-Myers Squibb using Pocock–Simon-type dynamic balancing algorithm to minimise imbalance in strata (study site, stage and PS). No indication of the proportion of patients recruited was provided. No placebo control
- Characteristics of the groups were similar, although for the number of disease sites, 53% in PAX+BSC arm had two sites versus 35% in BSC arm
- Blinding was not specified
- Two groups were directly compared. Statistical significance was indicated at $p < 0.05$ level. CIs were given for survival and time to progression. At least 144 patients were necessary to detect a 20% difference in survival at 1 year in the BSC arm, 18-month accrual and a minimum of 2 years of study time, with 80% power and alpha level of 0.05. Recruitment continued to 10% above required level (to allow for any non-evaluable patients). Survival curves were estimated using the Kaplan–Meier method on the entire randomised population and compared for statistical significance by a two-tailed log-rank test. Cox regression analysis for survival was performed, stratified by tumour stage and PS. The comparative analysis of QoL data was predetermined to extend until there was < 30% of the total randomised patients remaining available for QoL assessment. QoL treatment comparisons over time were assessed by longitudinal analysis for the median change from baseline for the psychological, physical and functional activity scores, and the questions of general activity and general QoL with the use of the Wei-Johnson test of stochastic ordering
- One patient in the PAX+BSC arm withdrew prior to chemotherapy. Three patients in the PAX+BSC arm were unevaluable for response assessment (1 never treated, 1 with non-lung cancer tumour history and 1 with no follow-up tumour evaluation)

General comments

- In total, 23 patients received off-protocol chemotherapy: 7 patients in the PAX+BSC arm received additional chemotherapy, and 16 patients in the BSC arm received non-PAX+BSC chemotherapy (14 received combination chemotherapy, and 2 received single-agent CDDP). These patients were included in the analysis
- Inclusion and exclusion criteria were clearly defined. Additional inclusion criteria included: adequate bone marrow function (absolute neutrophil count of $\geq 1.5 \times 10^9$/l and platelet count of $\geq 100 \times 10^9$/l); no history of prior or concomitant malignancy (except for curatively treated non-melanoma skin cancer or carcinoma *in situ* of the cervix), prior cytotoxic chemotherapy, requirement for urgent radiotherapy, symptomatic brain metastases, serious cardiac uncontrolled cardiac disease, myocardial infarction within the previous 6 months
- Previous radiotherapy exposure threshold was lowered 1 year into study (from 2 months to 4 weeks)
- Appropriate outcome measures were used
- Uncertain why only 95% CI was reported for 1-year survival (median survival at 1 year was not reported)
- *Conflicts of interest*: Bristol-Myers Squibb did the randomisation, and four of the authors were employed by them

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1 (used Pocock–Simon dynamic balancing algorithm to minimise imbalance in strata [study site, stage and PS])
Was the study described as double-blind?	Not possible due to comparison with BSC
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	1 patient in the intervention group withdrew

Reference and design	Intervention	Participants	Outcome measures
Gatzemeier *et al.*, 2000[34] (Europe) Phase III, multicentre randomised trial Supported by Bristol-Myers Squibb Jadad score: 2/5	***PAX+CDDP arm*** PAX 175 mg/m^2 (iv) and CDDP 80 mg/m^2 (iv) on day 1 of 21-day cycle ***CDDP arm*** CDDP 100 mg/m^2 (iv) on day 1 of 21-day cycle ***Length of treatment*** Minimum of 3 cycles (patients with stable disease received up to 6 cycles) ***Other interventions used*** Before PAX administration – dexamethasone, cimetidine and diphenhydramine For CDDP – dexamethasone, metoclopramide and lorazepam	*n* = 414 patients PAX+CDDP arm: *n* = 207 CDDP arm: *n* = 207 81% of patients were male 82% of patients had PS of 80–100 70% of patients had stage IV disease 96% of patients had measurable disease 45% of patients had adenocarcinoma ***Characteristics of target population*** Histologically confirmed stage IIIB or IV NSCLC (measurable or non-measurable disease) Aged between 18 and 75 years Median age: PAX+CDDP arm, 60 years (range, 33–75 years); CDDP arm, 60 years (range, 32–75 years) PS ≥ 60 (Karnofsky scale) Patients allowed to have received palliative radiotherapy to < 30% of marrow-bearing bones > 4 weeks before study entry See *General comments* below ***Setting*** Inpatient or outpatient (depending on usual practice of the study site)	***Primary end-point*** Survival: Calculated from the date of randomisation to death or censored on the last day known to be alive ***Secondary end-points*** Time to progression: Date of randomisation until the date progressive disease first reported Response: WHO criteria. Time of evaluation not reported (possible before each treatment cycle) QoL: EORTC QLQ-C30 and LC13 questionnaires. Administered at randomisation (baseline), within 3 days before each cycle, at off-study and after off-study therapy every 2 months until progression Safety: Based on laboratory tests and on clinical signs and symptoms experienced during treatment period. Time of evaluation not reported (possible before each treatment cycle) See *General comments* below ***Length of follow-up*** Not specified

Results

- *Survival*: Survival curves shown; both arms comparable. At the time of analysis (not reported), a total of 335 patients of 414 had died: PAX+CDDP arm, 168 patients (81%); CDDP arm, 167 patients (80%). Median survival time was 8.1 months (95% CI, 7.3 to 9.2 months; range, 0.13–25.3+ months) for the PAX+CDDP arm and 8.6 months (95% CI, 7.1 to 10.3 months; range, 0.43–26.3+ months) for the CDDP arm ($p = 0.862$). The hazard ratio of 0.981 supports the fact that survival was comparable but lower in the PAX+CDDP arm. One-year survival was 30% (95% CI, 24% to 36%) for the PAX+CDDP arm and 36% (95% CI, 29% to 42%) for the CDDP arm
- *Median time to progression*: PAX+CDDP arm, 4.1 months (95% CI, 3.3 to 4.4 months); CDDP arm, 2.7 months (95% CI, 2.3 to 3.2 months)
- *Response*: Complete response was observed in 2% of patients in the PAX+CDDP arm and 1% of patients in the CDDP arm. Partial response, 25% vs 17%; stable disease, 52% vs 48% progressive disease, 22% vs 35%; for PAX+CDDP and CDDP arms, respectively. Overall response rate was 24% (50 of 207 patients; 95% CI, 18% to 31%) for the PAX+CDDP arm and 16% (34 of 207 patients; 95% CI, 12% to 22%) for the CDDP arm ($p = 0.47$)
- *QoL:* For the symptom scales, nausea/vomiting ($p = 0.0003$), appetite loss ($p = 0.02$) and constipation ($p = 0.032$) were significantly worse in the CDDP arm. Hair loss ($p = 0.0001$) and peripheral neuropathy ($p = 0.0001$) were significantly worse in the PAX+CDDP arm. For the functioning scales, only physical functioning was observed to be improved in the PAX+CDDP arm ($p = 0.054$)
- *Haematological toxicity*: Grades 3/4: neutropenia (45% vs 17%; $p < 0.05$), thrombocytopenia (1% vs 2%), anaemia (10% vs 6%), febrile neutropenia (4% vs < 1%; $p < 0.05$), infections (4% vs 6%), for PAX+CDDP and CDDP arms, respectively
- *Non-haematological toxicity*: Grade 3: peripheral neuropathy (4% vs 1%), arthralgia/myalgia (5% vs 2%), ototoxicity (0% vs 1%), renal toxicity (< 1% vs 1%), grades 3/4 nausea/vomiting (12% vs 16%), for PAX+CDDP and CDDP arms, respectively
- *Dosage*: In total, 795 courses of CDDP and 8899 courses of PAX/CDDP were administered. Median dosage: 3 cycles (range, 1–9 cycles) of CDDP and 5 cycles (range, 1–8 cycles) of PAX/CDDP

Comments

Methodological comments

- Randomisation was stratified by institution, Karnofsky PS (60–70 vs 80–100) and stage of disease (IIIB vs IV). Randomisation was performed centrally using a Pocock-type minimisation procedure for stratified randomisation
- Prognostic factors, of age, gender, PS, stage of disease and histology, were equally distributed in both arms
- Blinding was not specified
- ITT analysis was used for survival, response rate and time to progression
- Sample size was calculated on a 1-year survival rate of 25% (i.e. median survival of 6 months) for the high-dose CDDP arm. Using two-sided alpha level of 5%, a total of 400 patients were to be randomised over 18 months and observed for 7 months to provide at least an 85% power to detect a 50% relative improvement in the 1-year survival rate of the PAX+CDDP arm compared with the CDDP arm (i.e. an absolute survival rate or difference of 12.5% in 1-year survival rate or a difference in median survival of 2.5 months). Kaplan–Meier estimates were used in the analysis of all time-to-event variables (survival, time to progression, duration of response). A 95% CI for the median time to event was computed using the method of Brookmeyer and Crowley. A log-rank test was stratified by the prognostic stratification factors for a comparison of survival and time to progression. QoL analysis was performed using the five functional scales, the global health status scale and the 20-symptom scales (recommended by EORTC). For each scale, the difference from baseline was compared between treatment arms with a null hypothesis of equal distributions per time point vs a stochastic ordering alternative using a non-parametric longitudinal Wei-Johnson test. Toxicity was evaluated considering the worst reported event per patient. For each toxicity, treatment arms were compared using Fischer's exact test for 2 x 2 tables, for occurrence of any toxicity and for occurrence of severe toxicity
- Of the 414 patients randomised, 6 patients never received study drugs (CDDP arm, 1 patient; PAX+CDDP arm, 5 patients)

continued

General comments

- Exclusion and eligibility criteria were clearly described. Additional inclusion criteria included: patients with asymptomatic brain involvement were eligible, provided that it was not the only disease site; adequate baseline bone marrow, hepatic and renal functions were required; no history of prior or concomitant malignancy, except for curatively treated non-melanoma skin cancer or *in situ* cervical cancer or other cancer cured by surgery alone with a disease-free survival longer than 5 years; no prior history of atrial or ventricular arrhythmias and/or history of congestive heart failure, even if medically controlled; no prior documented myocardial infarction or pre-existing motor or sensory neurological symptoms ≥ grade 2 (WHO criteria); no active infections or other serious underlying medical conditions)
- End-points of the trial were defined, although concern exists that time of evaluation was not clearly specified for either response rate or toxicity
- No information on baseline or treatment evaluation
- Concern exists that standardised safety criteria were not used by the study. Reporting of QoL data was inadequate
- *Conflicts of interest:* Trial supported by Bristol-Myers Squibb

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	0
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	Of the 414 patients randomised, 6 patients never received study drugs (CDDP arm, 1 patient; PAX+CDDP arm, 5 patients)

TOI, Trial Outcome Index; SE, standard error; NCIC-CTC, National Cancer Institute of Canada – Common Toxicity Criteria; LVEF, left ventricular ejection fraction

Appendix 8

Summary of evidence of effectiveness of vinorelbine in lung cancer

Reference and design	Intervention	Participants	Outcome measures
Baldini *et al.*, 1998[36] (Italy) Phase II, multicentre randomised study Funding: Not specified Jadad score: 3/5	***Group 1*** CDDP 80 mg/m^2 (iv) on day 1, VDS 3 mg/m^2 (iv) on days 1 and 15, with MITO 6 mg/m^2 (iv) on day 1 of every 28-day cycle ***Group 2*** CDDP 80 mg/m^2 (iv) on day 1, IFOS 3 mg/m^2 (iv) on day 1, with VNB 25 mg/m^2 (iv) on days 1 and 8 of every 21-day cycle ***Group 3*** CBDCA 350 mg/m^2 (iv) on day 1, with VNB 25 mg/m^2 (iv) on days 1 and 8 of every 28-day cycle ***Length of treatment*** Minimum of 2 cycles, unless rapid tumour progression was documented. Complete response, partial response and stable disease were treated with maximum of 6 courses ***Other interventions used*** Mesna 600 mg/m^2 before IFOS infusion and 1200 mg/m^2 orally 4 and 8 hours after IFOS, oral ciprofloxacin 1 g/dl and fluconazole 50 mg/d given for prophylaxis	*n* = 140 Group 1: *n* = 49 Group 2: *n* = 48 Group 3: *n* = 43 86% of patients were male 67% of patients had stage IV disease 85% of patients had PS of 0 or 1 ***Characteristics of target population*** Histologically or cytologically proven stage IIIB/IV NSCLC Aged ≤ 75 years Median age: Group 1, 62 years (range, 37–69 years); Group 2, 64 years (range, 45–73 years); Group 3, 61 years (range, 47–72 years) No prior chemotherapy Presence of bi-dimensionally measurable disease WHO PS ≤ 2 See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: WHO criteria, evaluated after 2 courses of therapy ***Secondary end-points*** Toxicity: WHO criteria, evaluated after each therapy cycle Survival: Calculated from date of randomisation See *General comments* below ***Length of follow-up*** Not specified (at least 1 year)

Results

- *Response*: No complete responses achieved. Overall response rate: Group 1, 14.3% (95% CI, 5.94% to 27.2%); Group 2, 16.7% (95% CI, 7.4% to 30.2%); Group 3, 14% (95% CI, 5.3% to 27.9%)
- *Survival*: There were 9 (18%) patients from Group 1, 7 (15%) patients from Group 2 and 7 (16%) patients from Group 3 alive at the end of 1 year (randomisation occurred from August 1993 to October 1994). Median overall survival: Group 1, 8.4 months; Group 2, 8.8 months; Group 3, 7.9 months. No significance levels were provided
- *Haematological toxicity*: Main side-effect was myelosuppression. Grades 3 and 4: leucopenia (10.2% and 4.1% vs 16.7% and 8.3% vs 9.3% and 9.3%), thrombocytopenia (4.1% and 2% vs 2.1% and 4.2% vs 7% and 0%), anaemia (6.1% and 0% vs 10.4% and 0% vs 2.3% and 2.3%), for Group 1, Group 2 and Group 3, respectively. Grade 4 nephrotoxicity was reported in 2.1% of patients in Group 2 only
- *Non-haematological toxicity*: Nausea and vomiting grade 3 only: Group 1, 8.2%; Group 2, 12.5%; Group 3, 2.3%
- Toxicity was evaluable in 128 patients; 20 patients (Group 1, 9 patients; Group 2, 6 patients; Group 3, 5 patients) required dose and schedule modifications due to toxicity. Three toxic deaths were reported: 2 deaths in Group 3 because of neutropenic fever and sepsis, and 1 death in Group 3 due to adynamic ileus

continued

Comments

Methodological comments

- Randomisation (centralised) was performed by telephoning the trial office at the National Institute of Cancer Research in Genoa, and patients were assigned to one of the three regimens, using a computer-generated list stratified according to centre. Not placebo-controlled
- Patient characteristics were very similar in all three treatment arms
- The study coordinator and an expert radiologist were blinded to treatment; they assessed the responses alongside the review committee
- All randomised patients were included in an ITT analysis (early deaths and early progressions were considered treatment failures). Response rates used ITT analysis. Due to study design, a formal comparison of all three regimens was not planned. Survival curves were plotted using Kaplan–Meier method. Toxicity was evaluable in 128 patients
- Simon's optimal two-stage design for Phase II clinical trials was used to calculate sample size and to minimise expected number of patients to be accrued in case of low-activity combination. Sample size was calculated assuming: alpha error of 0.05, beta error of 0.10; P0 (clinically uninteresting true response rate) and P1 (sufficiently promising true response rate), defined according to Simon, were set at 10% and 30%, respectively. In the first stage, 18 patients in each arm had to be randomised; if two or less responses were observed, the accrual had to be stopped; otherwise, 17 more patients had to be accrued. The drug combination was considered of interest if seven or more responses were observed out of 35 evaluable patients
- Of the 140 patients, 3 were not eligible: 2 patients from Group 1 (1 patient with SCLC and 1 with stage IIIB disease due to tracheal invasion) and 1 from Group 2 (brain metastases). Seven early deaths occurred (Group 1, 6 deaths; Group 2, 1 death). There were 19 patients (Group 1, 7 patients; Group 2, 9 patients; Group 3, 3 patients) who were not evaluable for response: 7 patients were not evaluable because of inadequate follow-up (mostly lack of confirmation of response after 4 weeks), 6 patients because of inadequate response documentation, and 6 patients refused treatment

General comments

- Inclusion and exclusion criteria were clearly defined. Additional inclusion criteria included: normal haematological function (Hb, WBC, platelet count); normal renal function; normal liver function; no active CNS disorder or brain metastases; no cardiovascular disease (cardiac failure, myocardial infarction within the previous 3 months, uncontrolled hypertension or arrhythmias); no concomitant neoplasm, other than *in situ* cervical carcinoma or cutaneous basal cell cancer. Patients in relapse after surgery were eligible, and patients previously treated with radiotherapy were eligible if they had other indicator lesions outside the irradiated area
- Interventions and outcome measures were clearly defined. Concern exists that, although duration of response was suggested to be a secondary end-point in the 'statistical analysis' section of the article, the results do not report any of the findings. Further concern exists that there were no significance levels provided for survival, response or toxicity end-points
- *Time of assessment:* Pretreatment evaluations included patient history and physical examination, WBC and chemistry profile, ECG, fibre-optic bronchoscopy, chest radiography, thoracic CT scan, abdominal CT scan or ultrasound. Bone scan or skeletal radiography and brain CT scan were performed only when clinically indicated. During treatment, WBCs, with differential and platelets, were performed weekly; a physical examination and chemistry profile were repeated before day 1 of each cycle. All responses were checked by the review committee, including the study coordinator and an expert radiologist, who were not aware of the treatment
- *Conflicts of interest:* None

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	3/140 patients were not eligible: 2/3 patients from Group 1 (1 patient with SCLC and 1 patient with stage IIIB disease due to tracheal invasion) and 1/3 from Group 2 (brain metastases) 7 early deaths (Group 1, 6 deaths; Group 2, 1 death) 6 patients refused treatment Patients not evaluable: 12 for toxicity and 19 for activity/efficacy

Reference and design	Intervention	Participants	Outcome measures
Colleoni *et al.*, 1997[37] (Italy) Phase II, randomised trial Funding: Not specified Jadad score: 1/5	***Group A*** CDDP 100 mg/m^2 (iv) on day 1, with MITO 8 mg/m^2 (iv) on day 1 and VNB 25 mg/m^2 (iv) on days 1 and 8, of a 21-day cycle ***Group B*** CBDCA 400 mg/m^2 (iv) on day 1, with VNB 25 mg/m^2 (iv) on days 1 and 8, of a 21-day cycle Treatments were discontinued and doses amended depending on the significance of side-effects ***Length of treatment*** Not prespecified. There were 98 cycles in Group A and 91 cycles in Group B, with a median of 4 cycles/patient. Dose reductions occurred in 27 cycles in Group A and 28 cycles in Group B ***Other interventions used*** Fluids 3000 cc, 5% dextrose solution, normal saline	*n* = 52 Group A: *n* = 26 Group B: *n* = 26 92% of patients were male 87% of patients had PS of 0 or 1 52% of patients had histology of squamous cell carcinoma 60% of patients had stage IIIB disease ***Characteristics of target population*** Histologically confirmed unresectable NSCLC TNM disease stage IIIB–IV Aged < 70 years Median age: Group A, 62 years (range, 42–72 years); Group B, 63 years (range, 47–71 years) Untreated with systemic chemotherapy or immunotherapy PS of 0–2 (ECOG scale) Measurable disease See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: WHO criteria, evaluated every 2 cycles ***Secondary end-points*** Side-effects (WHO criteria) Time to progression: Calculated from first day of treatment Survival: Calculated from first day of treatment See *General comments* below ***Length of follow-up*** Not clearly stated; appears limited to treatment period

Results

- *Response*: No complete response reported in either arm. Partial response, 11 patients (42%) vs 7 patients (27%); progressive disease, 6 patients (23%) vs 7 patients (27%); stable disease, 9 patients (35%) vs 12 patients (46%), for Group A and Group B, respectively. Overall response: Group A, 11/26 patients (42%; 95% CI, 23% to 63%); Group B, 7/26 patients (27%; 95% CI, 12% to 48%). Response according to stage of disease: Group A, 8/15 patients (53%) with stage IIIB and 3/11 patients (27%) with stage IV responded; Group B, 5/16 patients (31%) with stage IIIB and 2/10 patients (20%) with stage IV responded. Median duration of response: Group A, 4 months (range, 3–8 months); Group B, 4 months (range, 2–7 months)
- *Survival*: Median survival time: Group A, 9.9 months (range, 3–14 months); Group B, 8.8 months (range, 1–18 months). Median time to progression: Group A, 4.9 months (range, 3–9 months); Group B, 4.7 months (range, 1–9 months)
- *Haematological toxicity*: Grades 3 and 4: leucopenia (8% vs 12%), neutropenia (8% vs 15%), thrombocytopenia (0% vs 4%), anaemia (0% vs 8%), for Group A and Group B, respectively
- *Non-haematological toxicity*: Grades 3 and 4: infections (0% vs 0%), mucositis (0% vs 4%), phlebitis (0% vs 0%), vomiting (4% vs 0%), pain (0% vs 0%), fever (0% vs 0%), renal toxicity (4% vs 0%), neurotoxicity (0% vs 0%), for Group A and Group B, respectively

Comments

Methodological comments

- Patients were randomly allocated to either Group A or B
- Patients were well balanced between arms, according to age, gender, PS and disease stage
- Blinding was not specified
- Kaplan–Meier method was used to plot estimated survival and time to progression curves
- Given the objective of the study was to obtain 30% objective remissions, the Gehan two-step statistical approach was used. According to this approach, a total of 9 patients in the first step had to be studied in each arm, and if one response was seen, then a total of 25 patients per arm had to be studied
- No description of attrition/dropout rates was provided

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate renal function (serum creatinine < 1.2 mg/dl, blood urea nitrogen < 50 ml/dl) and liver function (bilirubin < 1.2 mg/dl, AST and ALT less than two times the upper limit of the norm); adequate bone marrow reserve (WBC > 4000/mm^3 and platelets > 100,000/mm^3); no severe concomitant diseases, active infections or cerebral metastasis
- Interventions were clearly defined. Outcome measures were clearly defined
- Complete response considered as a complete disappearance of all evident tumour, determined by two observations not less than 4 weeks apart; a partial response considered as a > 50% decrease in the cross-sectional areas of the measurable lesions. Pretreatment examination included: complete blood counts; biochemistry; chest X-ray; ECG; abdominal, thoracic and cerebral CT scan and isotope bone scan. Weekly assessment: complete blood count. Every 21 days (end of cycle): biochemistry tests
- *Conflicts of interest*: None specified

continued

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	0
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	Not stated

Reference and design	Intervention	Participants	Outcome measures
Colucci *et al.*, 1997[38] (Italy) Phase III, multicentre, randomised trial Funding: Not specified Jadad score: 2/5	***Group 1*** CDDP+VNB (step 1), IFOS+EPI (step 2). CDDP 100 mg/m^2 (iv) on day 1, with VNB 25 mg/m^2 (iv) on days 1 and 8 of every 21-day cycle for 3 cycles, then IFOS 2.5 g/m^2 (iv) with mesna on days 1 and 2, with EPI 100 mg/m^2 (iv) on day 1 of every 21-day cycle for 3 cycles ***Group 2*** IFOS+EPI (step 1), CDDP+VNB (step 2). IFOS 2.5 g/m^2 (iv) with mesna on days 1 and 2, with EPI 100 mg/m^2 (iv) on day 1 of every 21-day cycle for 3 cycles, then CDDP 100 mg/m^2 (iv) on day 1, with VNB 25 mg/m^2 (iv) on days 1 and 8 of every 21-day cycle for 3 cycles ***Length of treatment*** Six cycles ***Other interventions used*** Saline, magnesium sulphate and potassium chloride supplements, mannitol, mesna, anti-emetics (ondansetron or granisetron), methyl-prednisolone	*n* = 100 Group 1: *n* = 53 Group 2: *n* = 47 86% of patients were male 62% of patients had PS of 90–100 49% of patients had squamous cell carcinoma ***Characteristics of target population*** Histologically confirmed diagnosis of unresectable stage IIIA/IIIB or metastatic stage IV NSCLC Measurable disease according to the WHO criteria Aged ≤ 70 years Median age of Group 1: stage III, 64.5 years (range, 48–70 years); stage IV, 61 years (range, 41–70 years) Median age of Group 2: stage III, 64 years (range, 51–70 years); stage IV, 62 years (range, 28–69 years) PS ≥ 80 (Karnofsky scale) No previous chemotherapy and/or radiotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: WHO criteria, evaluated after 3 therapy cycles ***Secondary end-points*** Toxicity: WHO criteria, evaluated before each therapy cycle Time to progression: Calculated from the first day of chemotherapy until progressive disease was evidenced Survival: Calculated from the date of randomisation until the date of death or last follow-up See *General comments* below ***Length of follow-up*** Not specified

Results

- *Response (step 1)*: Complete response, 1 patient (2%) vs 1 patient (2%); partial response, 24 patients (45%) vs 9 patients (19%); stable disease, 9 patients (17%) vs 12 patients (26%); progressive disease, 19 patients (36%) vs 25 patients (53%), for Group 1 and Group 2, respectively. Overall response rate was significantly higher (p = 0.0112) for Group 1 (47%; 95% CI, 33% to 61%) than Group 2 (21%; 95% CI, 11% to 35%). Response rates were stratified according to stage: Stage III complete response, 1 patient (5%) vs 0 patients (0%); partial response, 11 patients (50%) vs 3 patients (14%); stable disease, 4 patients (18%) vs 6 patients (27%); progressive disease, 6 patients (27%) vs 13 patients (59%), for Group 1 and Group 2, respectively. Overall response rate was significantly higher (p = 0.0097) for Group 1 (55%) than for Group 2 (14%). Stage IV complete response, 0 patients (%) vs 1 patient (4%); partial response, 13 patients (42%) vs 6 patients (24%); stable disease, 5 patients (16%) vs 5 patients (20%); progressive disease, 13 patients (42%) vs 13 patients (52%), for Group 1 and Group 2, respectively. Overall response rate was not significantly higher (p = 0.4) for Group 1 (42%) in comparison with Group 2 (28%)
- *Response (step 2)*: Overall, 23 (43%) of patients in Group 1 and 19 (40%) of patients in Group 2 reached the second step of the study. Complete response, 0 patients (0%) vs 1 patient (5%); partial response, 0 patients (0%) vs 4 patients (21%); stable disease or progressive disease, 23 patients (100%) vs 14 patients (74%). Overall response rate was significantly higher (p = 0.037) for Group 2 (26%) than for Group 1 (0%). Completed full plan and underwent radiotherapy: Group 1, 6/23 patients (26%); Group 2, 4/14 patients (28%) (NS)
- *Median time to progression*: Group 1, 6 months; Group 2, 4 months (p = 0.665). Stage III: Group 1, 7 months; Group 2, 3 months (p = 0.049). Stage IV: Group 1, 6 months; Group 2, 5 months (p = 0.708)
- *Survival*: Median survival time: Group 1, 9 months; Group 2, 7 months (p = 0.328). Stage III: Group 1, 13 months; Group 2, 7 months (p = 0.03). Stage IV: Group 1, 7 months; Group 2, 7 months (p = 0.526). Median time to progression and overall survival of patients completing both steps 1 and 2 showed no statistically significant difference between the two arms. Median time to progression and overall survival of patients who achieved a major objective response and completed both treatments steps were also not statistically significant (p = 0.32)
- *Haematological toxicity*: Leucopenia was statistically more frequent during the administration of the IFOS+EPI regimen. Grades 3 and 4, step 1: WBC, 2 patients (4%) vs 10 (21%; p = 0.012); platelets, 0 patients (0%) vs 1 (2%); Hb, 5 patients (10%) vs 2 (4%), for Group 1 and Group 2, respectively. Grades 3 and 4, step 2: WBC, 1 patient (4%) vs 6 (31%; p = 0.034); platelets, 1 patient (4%) vs 1 (5%); Hb, 4 patients (17%) vs 3 (16%), for Group 1 and Group 2, respectively*
- *Non-haematological toxicity*: Grades 3 and 4, step 1: nausea/vomiting, 5 patients (10%) vs 7 (15%); stomatitis, 0 patients (0%) vs 2 (4%); diarrhoea, 1 patient (2%) vs 1 (2%), for Group 1 and Group 2, respectively. Grades 3 and 4, step 2: nausea/vomiting, 7 patients (31%) vs 2 (10%); stomatitis, 0 patients (0%) vs 2 (10%); diarrhoea, 0 patients (0%) vs 1 (5%), for Group 1 and Group 2, respectively*

continued

Comments

Methodological comments

- Patients were randomised to one of two treatment arms. Randomisation occurred centrally and was stratified according to stage (III vs IV) and PS (100–90 vs 80)
- Two treatment arms were well balanced in terms of age, PS, stage (III vs IV) and histological types, with no statistically significant difference between any subgroups. A 16% excess in patients with PS of 90–100 was observed for Group 2 over Group 1 in stage IV patients, but this figure was not statistically significant ($p = 0.58$). Squamous cell carcinomas represented the predominant histological type, but a 14% excess in the frequency of adenocarcinomas was observed for Group 2 over Group 1 in stage IV patients, but this was not statistically significant ($p = 0.567$)
- Blinding was not specified
- ITT analysis was used. Chi-squared test was applied to the contingency table to analyse if prognostic factors were well balanced between treatment arms 1 and 2. Logistic linear analysis was carried out to evaluate the effects of potential prognostic factors on response rate. The Kaplan–Meier product-limited method was used to generate time to progression and survival curves. Curve comparison was carried out by the log-rank test. Fisher's exact test used to test for statistical significance
- Sample size calculation was based on a 25% difference between treatments. Thus, 102 patients had to be randomised to detect a 25% difference in response rate or time to progression between two groups of patients at the significance level of $\alpha = 0.1$ with an 80% power ($\alpha = 0.8$)
- In Group 1, 57% of patients did not reach step 2; in Group 2, 60% did not reach step 2

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: life expectancy ≥ 3 months; absence of brain metastases; absence of second malignancies, with the exception of cutaneous basiloma or adequately treated *in situ* uterine carcinoma; WBC ≥ 4000/mm^3, platelets ≥ 120,000/mm^3, Hb ≥ 10 g %; adequate liver (serum bilirubin ≤ 1.2 mg %, transaminases less than twice their normal value) and renal functions (serum creatinine < 1.2 mg %, BUN ≤ 50 mg %, creatinine clearance ≤ 60 ml/minute); absence of uncontrolled severe cardiovascular, metabolic, neurological or infectious diseases
- Interventions were clearly defined. Outcome measures were clearly defined
- Concern exists that the study was not powered to detect difference in response rate (102 patients required, 100 patients randomised)
- Complete response defined as complete disappearance of all signs of disease for at least 4 weeks; partial response defined as ≥ 50% reduction in the sum of the products of the largest perpendicular diameters of measurable lesions for at least 4 weeks without the appearance of any new metastatic deposit or increase in size of any pre-existing lesion. Basal work-up included complete medical history and physical examination, complete blood cell counts and serum chemistries, standard 2p chest radiograph, abdominal sonogram, ^{99}Tc bone scan and ECG. All patients had CT scan of the thorax and upper abdomen. Before receiving high-dose EPI, all patients were given echocardiography with evaluation of the LVEF. Toxicity was assessed after every treatment cycle
- *Conflicts of interest*: None

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	***Group 1*** 30/53 patients (57%) did not reach step 2: 13/30 (43%) due to a drop in PS, 4/30 (13%) due to death, 6/30 (20%) due to locoregional treatment, 4/30 (13%) due to toxicity, 3/30 (10%) due to other reason ***Group 2*** 28/47 patients (60%) did not reach step 2: 16/28 (57%) due to drop in PS ($p = 0.43$, NS), 4/28 (14%) due to death, 2/28 (7%) due to locoregional treatment (Fisher's exact test, $p = 0.25$), 1/28 (4%) due to toxicity, 3/28 (11%) due to other reason, 2/28 (7%) due to too early (not defined)

[*] *For toxicity, step 1: Group 1,* n = *52; Group 2,* n = *47. For toxicity, step 2: Group 1,* n = *19; Group 2,* n = *23*

Reference and design	Intervention	Participants	Outcome measures
Comella *et al*., 1996[39] (Italy) Phase III, multicentre, randomised trial Funding: Not specified Jadad score: 3/5	***Group A*** CDDP 40 mg/m^2 (iv), with VP-16 100 mg/m^2 (iv) on days 1–3, for every 28-day cycle ***Group B*** CBDCA 250 mg/m^2 (iv) on day 1, with CDDP 30 mg/m^2 on days 2 and 3, with VP-16 100 mg/m^2 (iv) on days 1–3, with VNB 30 mg/m^2 on day 1, for every 28-day cycle ***Length of treatment*** Six cycles ***Other interventions used*** Saline, potassium chloride, antiemetic prophylaxis (anti-HT$_3$ receptors plus dexamethasone)	*n* = 112 (105 eligible) Group A: *n* = 57 (53 eligible) Group B: *n* = 55 (52 eligible) Of 105 patients eligible: 89% of patients were male 57% of patients had PS of 0 or 1 54% of patients had stage IV disease 53% of patients had squamous cell carcinoma 36% of patients had adenocarcinoma ***Characteristics of target population*** Histologically or cytologically proven diagnosis of NSCLC Advanced measurable disease (stage IIIB or IV) Aged ≤ 75 years Median age: Group A, 59.5 years (range, 35–72 years); Group B, 60.5 years (range, 40–73 years) PS ≤ 2 (ECOG criteria) No previous chemotherapy Life expectancy ≥ 3 months See *General comments* below ***Setting*** Not specified	***Primary end-point*** Objective response rate: WHO criteria, evaluated after 3 therapy cycles ***Secondary end-points*** Toxicity: WHO criteria, evaluated each week Survival: Evaluation not specified See *General comments* below ***Length of follow-up*** Median potential follow-up: Group A, 86 weeks; Group B, 82 weeks

Results

- At the end of December 1995, a total of 89 deaths had occurred, and 8 patients in each arm were still alive (randomisation occurred between March 1993 and June 1995). Based on the lower response rate in the experimental arm compared with the standard treatment arm, the accrual was stopped and the null hypothesis accepted
- *Response rate*: Group A: complete response, 0 patients (0%); partial response, 15 patients (28%); progressive disease, 12 patients (23%); stable disease, 14 patients (26%). Overall response for Group A: 28% (95% CI, 17% to 42%). Group B: complete response, 1 patient (2%); partial response, 12 patients (23%); progressive disease, 12 patients (23%); stable disease, 11 patients (21%). Overall response for Group B: 25% (95% CI, 13% to 37%). The difference was not statistically significant. Mantel-Haenszel chi-squared test was performed, which showed that treatment did not significantly affect the probability of response (p = 0.7), even after adjustment for histology (squamous vs others)
- On multiple logistic analysis, the treatment failed to affect the probability of response rate significantly. Among pretreatment features: stage IIIB (regression coefficient ± SE, 0.98 ± 0.47; relative risk, 2.7 (95% CI, 1.0 to 6.8; p = 0.037) and PS 0–1 (regression coefficient ± SE, 1.04 ± 0.52; relative risk, 2.8 (95% CI, 1.0 to 8.0; p = 0.046) were the only parameters independently predictive of a higher response rate
- *Survival curves*: Group A curve was above that of Group B until approximately 30 weeks, after which Group B curve was above. Median survival time: Group A, 31 weeks; Group B, 27 weeks. The difference was not statistically significant. Using multivariate Cox analysis, the type of treatment failed to show a significant impact on survival, while the outcome of the patients was significantly affected by stage and PS
- *Survival curves according to PS*: Group A with PS 0–1 curve was above all other curves at all points, after 26 weeks of survival. Group B with PS 0–1 curve was above all others until approximately 26 weeks of survival. Group A with PS 2 curve was above Group B with PS 2 curve at all points
- *Haematological toxicity*: Myelosuppression was the most frequent and limiting side-effect. Grades 3 and 4: neutropenia (5% and 1% vs 12% and 5%), thrombocytopenia (1% and 0% vs 5% and 4%), anaemia (5% and 0% vs 4% and 1%), for Group A and Group B respectively. Both neutropenia (p = 0.000006) and thrombocytopenia (p = 0.000007) were significantly more frequent in Group B than Group A, although significance was based on any grade in the two arms. Persistent neutropenia caused 1 patient in Group B to suspend therapy. Platelet transfusion was required by 2 patients in Group B, but no clinically significant haemorrhagic episodes were encountered
- *Non-haematological toxicity*: Grade 4 non-haematological toxicity did not occur, except for nausea/vomiting in Group A (4%) vs Group B (4%). Grade 3: nausea/vomiting (6% vs 8%), nephrotoxicity (1% vs 2%), diarrhoea (0% vs 1%), ototoxicity (1% vs 0%), for Group A and Group B, respectively (significance levels not provided)

continued

Comments

Methodological comments

- A centralised telephone call procedure was used to randomly assign individuals to either treatment arm based on a computer-generated list, stratified according to stage (IIIB vs IV) and PS (0–1 vs 2)
- No significant differences between the two treatment arms were found with respect to age, sex, stage, PS or weight loss. A significant imbalance (Fisher's exact test, $p = 0.02$) was observed in distribution of histological subtypes (more patients in Group B had non-squamous histology)
- Two radiologists and two oncologists blinded to treatment were responsible for assessing objective responses according to WHO criteria
- ITT analysis was used (112 patients included). In total, 105 patients were included in the response analysis. Fisher's exact test was used for comparison between group frequencies. Main pretreatment variables (PS, stage, histology, weight loss and age) together with treatment type were included in logistic linear model to determine the effect of treatment on response rate when adjusted for the main prognostic features. Survival curves were plotted using the Kaplan product-limiting method. Comparisons were made using the log-rank test. The Cox proportional hazard model was used to evaluate the effect of treatment on survival after adjusting for the main pretreatment variables; adjusted relative risks were calculated as antilogarithms of the regression coefficient. All analyses were performed using the Systat software package
- Sample size was established by using a two-stage optimal design for Phase III trials with binary response. Setting the errors alpha and beta at 5% and 20%, respectively, 52 patients for each arm were required to be randomised in the first stage. If actual response rate did not exceed that of the control arm by 3%, accrual was stopped and the experimental combination was rejected
- Seven patients were considered ineligible because they had stage IIIA disease (2 patients) or did not meet the haematological (3 patients) or PS (2 patients) requirements. Also, there were 5 toxic deaths, 4 patients withdrew because of toxicity, 3 were lost to follow-up, and 1 refused to continue with treatment

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate bone marrow reserve (WBC, platelets, Hb); normal liver and renal function; no congestive heart failure or angina; no serious arrhythmia; no recent myocardial infarction, uncontrolled infections or metabolic diseases
- Interventions and outcomes were clearly defined
- Concern exists that the baseline characteristics of the patients were only for eligible patients and not for all patients randomised to treatment
- At entry, a complete physical and clinical examination as well as medical history were performed, including assessment of weight loss and PS, and complete blood cell count and serum chemistries. Extent of disease was evaluated by chest radiograph, CT scan of thorax and upper abdomen, liver ultrasound scan and bone scintigraphy. Brain CT scan was performed to rule out cerebral metastases. Bone radiography was limited to suspicious areas. Physical examination and blood count were performed weekly. During each cycle, patients underwent physical examination, together with blood count and chemistry. At restaging, evaluation of all measurable lesions was performed using the same procedures employed before the beginning of therapy
- *Conflicts of interest*: None specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	7/112 patients (6%) were considered ineligible: 2/7 (29%) had stage IIIA disease, 3/7 (43%) did not meet haematological requirements, 2/7 (29%) did not meet PS requirements Also, 5/112 (4%) toxic deaths occurred, 4/112 patients (4%) withdrew due to toxicity, 3/112 (3%) were lost to follow-up, 1/112 (1%) refused to continue with treatment

Reference and design	Intervention	Participants	Outcome measures
Crawford *et al.*, 1996[40] (USA) Phase III, multicentre, randomised trial Funding: Glaxo Wellcome Jadad score: 3/5	***VNB arm*** VNB 30 mg/m^2 (iv) administered weekly ***5FU+LV arm*** LV 20 mg/m^2 (iv), followed by 5FU 425 mg/m^2 for 5 consecutive days every 4 weeks ***Length of treatment*** As long as disease was stable or responding ***Other interventions used*** Saline, 5% dextrose, haematopoietic growth factors; all patients received full supportive care (blood product transfusions, antibiotics, antiemetics, anti-diarrhoeals, analgesics as appropriate)	*n* = 216 VNB arm: *n* = 143 5FU+LV arm: *n* = 68 71% of patients were male in VNB and 5FU+LV arms PS ≥ 80 in 79% of VNB arm and 72% of 5FU+LV arm 48% of patients in VNB arm and 41% in 5FU+LV arm had adenocarcinoma ***Characteristics of target population*** Histologically or cytologically confirmed stage IV NSCLC Aged ≥ 18 years Median age: VNB arm, 61 years (range, 32–79 years); 5FU+LV arm, 61 years (range, 41–83 years) PS ≥ 70 (Karnofsky scale) Life expectancy ≥ 12 weeks No previous chemotherapy or biological therapy Adequate bone marrow, renal and hepatic function ≥ 2 weeks since prior surgery and 3 weeks since radiation therapy to the pelvis, long bone or spine See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Defined as the interval from the first day of drug treatment until the date of death ***Secondary end-points*** QoL: Physician assessment (PS and weekly symptom assessment) and patient assessment (SWOG questionnaire). Evaluated at baseline, before examination or before each therapy cycle Time to treatment failure: Calculated from the first day of treatment until the date of death Objective response rates: SWOG criteria, evaluated before each therapy cycle Duration of objective response: Calculated from the first day of treatment until the date of death Toxicity: Graded according to modified NCI criteria. Evaluated weekly Serial measurement of pulmonary function: Evaluated before each therapy cycle See *General comments* below ***Length of follow-up*** Not specified

Results

- *Survival*: Survival curves shown; VNB curve above 5FU+LV curve at all points until survival extended beyond day 400. Estimated median survival: VNB arm, 30 weeks with 25% alive at 1 year; 5FU+LV arm, 22 weeks with 16% alive at 1 year; (log-rank test, $p = 0.03$; Cox model, $p = 0.06$). Prognostic factors identified as significant in Cox proportional hazards model included metastatic disease to bone ($p = 0.002$), Karnofsky PS ($p = 0.003$), weight loss ($p = 0.012$) and pretreatment lactate dehydrogenase value > 350 U/l ($p = 0.021$). Non-significant prognostic factors included histology ($p = 0.127$), number of metastatic sites ($p = 0.837$), prior surgery ($p = 0.464$), presence of measurable vs non-measurable disease ($p = 0.385$) and treatment arm ($p = 0.062$)
- *QoL*: Analysis of results from QoL data indicated that patients treated with VNB stayed on study for longer and completed more questionnaires than 5FU+LV-treated patients; for patients who stayed on treatment, minimal changes in QoL scores were observed overall for each dimension for the two treatments (no data provided). There was a trend toward improved symptom distress scores for patients treated with VNB compared with 5FU+LV (no data provided). Physical functioning was better for 5FU+LV recipients compared with those who received VNB. Some improvements in mean global QoL scores were noted for both groups (no data provided). Physicians were instructed to query patients on cancer-related symptoms. No treatment differences were observed in relation to relief of cancer symptoms. During the first 8 weeks of treatment, 42 of 118 patients (36%) in VNB arm and 26 of 58 patients (45%) in 5FU+LV arm showed improvement in all baseline cancer symptoms
- *Response*: Two responders (no complete response reported, no further information provided) were observed in the first 25 patients recruited to 5FU+LV arm. ITT analysis (study not powered to detect differences in secondary end-points) of response rate was 12% for the VNB arm and 3% for the 5FU+LV arm (NS)
- *Median duration of response*: VNB arm, 23 weeks (8 weeks to > 52 weeks); 5FU+LV arm, 13 weeks and 25 weeks for 2 patients that responded
- *Overall time to treatment failure curve*: VNB curve was above 5FU+LV curve at all points. Median overall time to treatment failure: VNB arm, 10 weeks; 5FU+LV arm, 8 weeks (log-rank test, $p = 0.017$; Cox model, $p = 0.02$)
- *Haematological toxicity*: Principal haematological toxicity of VNB arm was granulocytopenia. Grades 3–4: granulocytopenia (54% vs 24%), granulocytopenic infections (7% vs 6%), thrombocytopenia (0% vs 2%), anaemia (1% vs 2%), for VNB and 5FU+LV arms, respectively
- *Non-haematological toxicity*: No Grade 4 non-haematological toxicities were reported for VNB arm. Grades 3–4: nausea (1% vs 2%), asthenia (5% vs 6%), stomatitis (0% vs 18%), diarrhoea (1% vs 7%), alopecia (1% vs 3%), for VNB and 5FU+LV arms, respectively

continued

Comments

Methodological comments

- Randomisation was based on 2:1 ratio of VNB to 5FU+LV (2:1 randomisation conducted so that at least 50 patients were enrolled into the VNB arm to ensure adequate clinical experience with VNB in multi-institutional setting). Randomisation was centralised, and treatment arms were computer generated in blocks of 6 patients: 4 patients to receive VNB and 2 patients to receive 5FU+LV. Not placebo controlled. Patients were stratified according to centre and presence of measurable and non-measurable disease at screening
- The two groups were comparable with respect to baseline prognostic factors, such as age, PS, weight loss, histology and gender. The 5FU+LV arm had a higher percentage of patients who received prior surgery (41% vs 27%, $p < 0.05$). The 5FU+LV arm also had higher percentage of patients with two or more sites of metastatic disease (66% vs 49%, $p < 0.05$)
- Blinding was not specified
- ITT analysis was used for all patients receiving study medication, but only patients with measurable disease (VNB arm, 126 patients; 5FU+LV arm, 58 patients) were included in response evaluation. Cox's proportional hazards model was used to compare survival data between the two groups. Kaplan–Meier curves were used to display data. QoL was analysed using repeated-measures of analysis of variance. Cancer-related symptom data were summarised over time, and logistic regression was used to compare treatment groups. Response rates were compared using Fisher's exact test. Time to treatment failure was compared using the log-rank test and Cox's proportional hazards model. CIs were not provided
- An estimated total of 150 patients were to be entered into the study with 2:1 randomisation of VNB to 5FU+LV. Statistical power of 0.80 was provided by sample size to detect a 12-week difference between the two treatment groups in median survival times. The study was not powered to detect differences in secondary end-points
- Five randomised patients (VNB arm, 1 patient; 5FU+LV arm, 4 patients) never received treatment (211 eligible patients: VNB arm, 143 patients; 5FU+LV arm, 68 patients). Two of the 5 patients died before the first dose, and the 3 remaining patients were lost to follow-up evaluation

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: previous radiotherapy was permitted if assessable disease was outside radiation portal; patients with measurable or non-measurable assessable disease were eligible; patients with measurable disease had to have bi-dimensionally measurable disease on CT scan, MRI, X-ray or physical examination; patients with only non-measurable assessable disease were required to have cancer-related symptoms; pleural effusions, ascites and laboratory parameters were not acceptable as the only evidence of assessable disease; not medically unstable; no CNS metastases or history of other malignancy, other than basal cell skin carcinoma or carcinoma *in situ* of the cervix; not presenting with clinically significant peripheral neuropathy not attributable to cancer
- Interventions and outcome measures were clearly defined. Tumour response rates: complete response was defined as the complete disappearance of all objective disease; partial response was defined as reduction in size of all measurable tumour areas of at least 50% from baseline without the appearance of any new disease, and no individual tumour increase of greater than 50% in the product of the bi-dimensional measurements. Complete response and partial response needed to be confirmed by a second evaluation at least 4 weeks later. Concern exists regarding the QoL measures used. Modified SWOG questionnaire was not a standardised instrument and, furthermore, was developed for use in prostrate and colon cancer, not lung cancer trials
- Concern exists in regards to the reporting of QoL results, with no data presented. Serial measurement of pulmonary function was stated as a secondary end-point, although no results were reported
- Assessment was performed weekly for toxicity, symptoms and adverse events, and before each treatment for blood counts, hepatic function and neurological toxicity. Evaluation was also performed 4 weeks after initial diagnosis of complete response or partial response. Baseline data were collected no more than 2 weeks before randomisation, including physical examination, history, PS and laboratory assessments. QoL questionnaire was completed before randomisation, before examination by physician or before patients received treatment. All assessments were completed when patients left study
- *Conflicts of interest*: Supported by Glaxo Wellcome

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	5/216 patients (2.3%) never received treatment: 1/5 (20%) were VNB patients, 4/5 (80%) were 5FU+LV patients. Reasons: 2/5 patients (40%) died before first dose, 3/5 (60%) were lost to follow-up evaluation (affiliation to arm not reported)

Reference and design	Intervention	Participants	Outcome measures
Depierre *et al.*, 1994[41] (France) Phase III, multicentre randomised trial Funding: Pierre Fabre Ltd Jadad score: 3/5	***VNB arm*** VNB 30 mg/m^2 (iv) per week VNB+CDDP arm VNB 30 mg/m^2 (iv) per week, with CDDP 80 mg/m^2 (iv) every 3 weeks ***Length of treatment*** Minimum of 6 weeks; treatment continued until progression. Duration of treatment of patients in complete remission was 27 weeks ***Other interventions used*** 5% dextrose solution, normal saline, furosemide, methylprednisolone, metoclopramide	*n* = 240 VNB arm: *n* = 119 VNB+CDDP arm: *n* = 121 86% of patients were male 55% of patients had stage IV disease 71% of patients had PS of 0 or 1 53% of patients had squamous cell carcinoma ***Characteristics of target population*** Histologically proven NSCLC Stage IIIA/B or IV disease Aged < 75 years Median age: VNB arm, 58.8 years; VNB+CDDP arm, 59.2 years No previous chemotherapy or radiotherapy WHO PS < 3 Stage III patients were to be unacceptable for surgery or radiotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response to treatment: WHO criteria, assessed after 9 weeks of treatment ***Secondary end-points*** Survival: Defined as the interval from the date of randomisation until the date of death, or last follow-up Time to progression: Defined as the interval from the date of randomisation until the date of progression (progression evaluated early if patients received at least 6 weeks therapy) Toxicity: WHO criteria, FBC repeated weekly, other laboratory parameters and chest radiographs repeated every 3 weeks ***Length of follow-up*** Complete restaging of patients after 9 weeks of treatment and subsequently every 3 months See *General comments* below

Results

- *Response*: There were 231 eligible patients (VNB arm, 115 patients; VNB+CDDP arm, 116 patients). Complete response, 1 vs 2 patients; partial response, 17 patients (2%) vs 48 patients (41%); stable disease, 28 patients (24%) vs 35 patients (30%); progressive disease, 58 patients (50%) vs 19 patients (16%), for VNB and VNB+CDDP arms, respectively. Overall response was 43% (95% CI, 34% to 52%) for the VNB+CDDP arm, which was significantly higher (p = 0.0001) than 16% (95% CI, 9% to 22%) for the VNB arm. The difference remained significant (p = 0.00005) when treatment effect (response vs no response) was adjusted based on prognostic variables (i.e. histology type, staging, PS and neutrophil counts)
- *Survival*: On the reference date (1 January 1992), 181 patients had died, 91 in the VNB arm and 90 in the VNB+CDDP arm (randomisation occurred between 1 October 1989 and 19 August 1991). Survival curves shown; curves intersect at approximately 50 days and then again at approximately 200 days. Median survival: VNB arm, 32 weeks; VNB+CDDP arm, 33 weeks (log-rank test, p = 0.48). Treatment comparison for overall survival (adjusted for histology type, stages, PS and polymorphonuclear neutrophil count) showed no modification of the risk of death in either of the estimated groups (score test, p = 0.34,). When survival was calculated in patients who were evaluable for response, the difference remained non-significant. Analysis of survival curves in all treated patients showed that obtaining a partial response was predictive of survival in both arms, with a median survival of 23 weeks for progression, 30 weeks for no change (in size of measurable lesions without appearance of new lesions) and 38 weeks for partial response (log-rank test, p = 0.0001). For a given response level, survival rates were higher in the VNB group (survival rates at 12 months for partial response, stable disease and progressive disease were, respectively, 63%, 24% and 7% in the VNB arm and 57%, 19% and 0% in the VNB+CDDP arm)
- *Time to progression*: Of the 231 eligible patients, 181 (77%; 97 patients in the VNB group and 84 patients in the VNB+CDDP group) showed evidence of progression from reference date. Median time to progression: VNB arm, 10 weeks; VNB+CDDP arm, 20 weeks (log-rank test, p = 0.0001). Not progressed at 6 months: VNB arm, 17% (95% CI, 10% to 24%); VNB+CDDP arm, 42% (95% CI, 32% to 52%)
- *Number of patients with partial response lasting 9 months or more*: VNB arm, 6 partial responders (duration, 9, 9, 9, 9, 10 and 17 months); VNB+CDDP arm, 7 partial responders (duration, 9, 11, 11, 12, 13, 13 and 14 months)
- *Haematological toxicity*: Evaluated in all 240 patients; WHO grades were not reported. Thrombocytopenia < 50,000 (0% vs 0%), leucopenia < 1000 (10% vs 23.1%), granulocytopenia < 500 (34.4% vs 65.3%), for VNB and VNB+CDDP arms, respectively
- *Non-haematological toxicity*: Grade 3–4 nausea/vomiting (5%, vs 23.1%), grade 2–3 renal toxicity (5.0% vs 29.7%), grade 3–4 constipation (5.8% vs 5.8%), grade 3–4 neurological events (1.7% vs 7.4%), grade 3–4 asthenia (11.7% vs 27.3%), grade 2 alopecia (5.9% vs 10.7%), for VNB and VNB+CDDP arms, respectively

continued

Comments

Methodological comments

- Randomisation occurred through a centralised blind telephone assignment procedure, with centre and stage stratification
- *Comparability of treatment groups*: The two groups were well balanced for PS, age, staging and histology, and there was no significant difference between the arms. However, for gender there was an imbalance, with a higher percentage of women in the VNB arm; this difference was statistically significant ($p = 0.03$)
- Treatment administration was not blind. Initial response, confirmation of response and date of progression were validated by a 'blinded' multidisciplinary group
- *Method of data analysis*: Survival analysis was made based on ITT analysis. Statistical analysis was based on findings at the reference date (1 January 1992). At this date, 188 deaths were effectively observed, with no patients lost to follow-up. Resulting power was 85%, not 90%. Treatment comparison used two-tailed tests of significance. Survival analysis was based on Kaplan–Meier's estimate, the log-rank test and Cox's regression model. Response analysis was based on the chi-squared test and a logistic regression model for adjustment on prognostic covariates. Calculations were based on the first 18 weeks of therapy. SAS and BMDP software packages were used for statistical analysis
- *Sample size/statistical power*: Estimation of sample size was based on the method described by George and Desu, with an expected annual accrual rate of 100 patients, a type I error α of 0.05, a type II error β of 0.10 for a one-sided test and an assumption of treatment benefit given by an increase in 1-year survival from 20% with VNB alone to 35% with the VNB+CDDP combination. A total of 188 deaths had to be observed before fixed-sample design statistical analysis would be appropriate
- Nine patients (3.8%) were considered ineligible (4 patients in VNB arm and 5 patients in VNB+CDDP arm) because of the association of current or previous other malignancy (4 patients), severe cardiac failure (1 patient), a PS > 2 (1 patient), initial brain metastases (1 patient), radiotherapy before randomisation (1 patient) or absence of any evaluable lesion (1 patient). There were 23/231 patients (10%) who were not evaluable for response: 11 (48%) were from the VNB arm and 12 (52%) were from the VNB+CDDP arm. Reasons were refusal of treatment (4 patients), early death due to intercurrent disorder (9 patients), non-fatal intercurrent illness (2 patients) or dropout because of toxicity (8 patients, including 4 toxic deaths)

General comments

- Eligibility and exclusion criteria were clearly described: at least one measurable lesion; bone lesion or pleural effusion were not acceptable as target lesions; normal hepatic function; normal renal function; no brain metastases or previous or concomitant malignancy, except basal cell skin carcinoma; no neurological or auditory history, or the existence of a non-controlled bacterial infection
- Interventions were clearly defined. Outcome measures were clearly defined
- First response assessment was performed 9 weeks after the beginning of treatment. Pretreatment evaluation included full physical examination, chest radiograph, lung CT scan, bronchofibroscopy, liver ultrasonography, adrenals and brain CT scan, isotopic bone scan, complete blood count, serum creatinine, transaminases, alkaline phosphatase and bilirubin. Follow-up evaluation: weekly, complete blood count; every 3 weeks, laboratory parameters and chest radiographs; after 9 weeks of treatment, complete restaging of patients; every 3 months, restaging of patients
- Concern exists that haematological toxicities were not reported by WHO severity grade
- *Conflicts of interest*: Supported by Pierre Fabre Oncologie (Boulogne, France)

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	9 patients (3.8%) were considered ineligible: 4/9 in VNB arm and 5/9 in VNB+CDDP arm. Reasons: association of current or previous other malignancy (4 patients), severe cardiac failure (1 patient), PS > 2 (1 patient), initial brain metastases (1 patient), radiotherapy before randomisation (1 patient) or absence of any evaluable lesion (1 patient) 23/231 patients (10%) were not evaluable for response: 11/23 (48%) from VNB arm and 12/23 (52%) from VNB+CDDP arm. Reasons: refusal of treatment (4 patients), early death due to intercurrent disorder (9 patients), non-fatal intercurrent illness (2 patients) or dropout because of toxicity (8 patients, including 4 toxic deaths)

Reference and design	Intervention	Participants	Outcome measures
Furuse *et al.*, 1996[42] (Japan) Phase III, crossover, multicentre, randomised trial Funding: Kyowa Hakko Company Jadad score: 3/5	***VNB arm*** VNB 25 mg/m^2 (iv) weekly. Patients not responding after 4 cycles of monotherapy were switched to VDS+ CDDP: VDS 3 mg/m^2 (iv) on days 1, 8 and 15, with CDDP 80 mg/m^2 (iv) on day 1, repeated every 4 weeks for at least 2 cycles. One administration is 1 cycle ***VDS arm*** VDS 3 mg/m^2 (iv) weekly. Patients not responding after 4 cycles of monotherapy were switched to VNB+ CDDP: VNB 20 mg/m^2 (iv) on days 1, 8 and 15, with CDDP 80 mg/m^2 (iv) on day 1, repeated every 4 weeks for at least 2 cycles. One administration is 1 cycle ***Length of treatment*** At least 4 cycles of monotherapy. If no response or progressive disease, patients were placed on combination therapy for at least 2 cycles ***Other interventions used*** Normal saline	*n* = 204 VNB arm: *n* =103 VDS arm: *n* =101 70% of patients were male 81% of patients had PS between 0 and 1 57% of patients had stage IV disease 61% of patients had adenocarcinoma ***Characteristics of target population*** Histologically or cytologically proven diagnosis of NSCLC Stages IIIB or IV (IUAC classification) disease Aged ≤ 77 years Median age: VNB arm, 67 years (range, 36–77 years); VDS arm, 68 years (range, 28–77 years) PS of 0, 1 or 2 (ECOG criteria) No prior treatment No symptoms of brain metastases No disseminated bone metastases requiring radiotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response rate: WHO criteria, evaluated each week Toxicity: WHO criteria, evaluated each week ***Secondary end-points*** Survival: Calculated from date of randomisation until date of death or last follow-up Median response: Calculated from allocation date to analysis See *General comments* below ***Length of follow-up*** Median duration of follow-up was 28.8 months, ranging from 5.7 months to 37.7 months

Results

- On 28 February 1995, 19 patients (9%) of population were still alive, and 2 patients in the VDS arm were lost to follow-up (randomisation between November 1991 and February 1993)
- *Response*: No complete response was observed in either arm. Partial response, 32 patients (31.1%; 95% CI, 22.3% to 40.8%) vs 9 patients (8.9%; 95% CI, 4.2% to 16.2%); progressive disease, 13 patients (13%) vs 26 patients (26%); stable disease, 56 patients (54%) vs 63 patients (62%), for VNB and VDS arms, respectively. Two patients in the VNB arm and 3 patients in the VDS arm had no assessable lesions. Median response duration: VNB arm, 18.5+ weeks (range, 7.9–107.3+ weeks); VDS arm, 11.7+ weeks (range, 6.0–35.0+ weeks). Cox regression analysis of response rates (univariate analyses): gender (male vs female, 13.3:36.1; p = 0.00042), age (≥ 65 vs ≤ 64 years, 19.5:21.0; p = 0.937), PS (0–1 vs 2, 18.1:28.9; p = 0.199), stage (IIB vs IV, 18.4:21.4; p = 0.728) and histology (squamous vs non-squamous, 11.3:23.9; p = 0.060). Gender was significantly related to response rate due to treatment (p = 0.00042). Results of multivariate analysis showed gender was an important prognostic factor for response (p = 0.0011)
- *Failure to respond*: In the VNB arm, 71 patients failed to respond; of these non-responders, 33 (46.5%) subsequently received VDS+CDDP, 9 underwent palliative radiotherapy and 7 received other chemotherapies; of the 16 patients receiving palliative radiotherapy or other chemotherapies, 3 had partial response; 22 patients (21.4%) refused further therapy. In the VDS arm, 92 patients failed to respond; of these non-responders, 49 (53.3%) subsequently received VNB+CDDP, 6 underwent palliative radiotherapy, and 13 received other chemotherapies; of the 19 patients receiving palliative radiotherapy or other chemotherapies, 3 had partial response; 24 patients (23.8%) refused further therapy
- *Overall objective response rate to combination therapy*: With VNB+CDDP, 13/49 patients (26.5%; 95% CI, 15% to 40.7%) had a partial response, 26 had no change, 6 had progressive disease, and 4 had no assessable lesions. With VDS+CDDP, 0/33 patients had partial response, 23 had no change, 7 had progressive disease, and 3 had no assessable lesions. Median duration of response of combination therapy: with VNB+CDDP, 25+ weeks (range, 14–45+ weeks)
- *Survival*: Survival curves shown; VNB curve above VDS curve for up to 500 days, after which both curves were relatively equal. Median survival time: VNB arm, 52.4 weeks; VDS arm, 43.6 weeks; difference was not statistically significant (Wilcoxon test, p = 0.3839)
- *Haematological toxicity – monotherapy*: Grades 3 and 4: leucocytopenia, 55.3% (95% CI, 45.7% to 64.9%) vs 48.5% (95% CI, 38.5% to 58.7%; NS); anaemia, 16.5% vs 6.9% (p = 0.014), for VNB and VDS arms, respectively
- *Non-haematological toxicity – monotherapy*: Nausea/vomiting, 1.9% vs 0% (NS); alopecia, 0% vs 1% (p = 0.001); peripheral neurotoxicity, 0% vs 2.9% (p = 0.002); constipation, 2.9% vs 3% (NS); local cutaneous reaction, 1% vs 0% (p = 0.012); fever, 0% vs 0% (NS); pulmonary events, 0% vs 1% (NS), for VNB and VDS arms, respectively
- *Haematological toxicity – combination therapy*: Leucocytopenia, 55.1% vs 60.6% (NS); anaemia, 36.7% vs 27.3% (NS), for VNB+CDDP and VDS+CDDP arms, respectively
- *Non-haematological toxicity – combination therapy*: Nausea/vomiting, 12.2% vs 6.1% (NS); alopecia 0% vs 3% (p = 0.006); peripheral neurotoxicity, 0% vs 0% (NS); constipation, 12.2% vs 12.1% (NS); local cutaneous reaction, 0% vs 0% (NS); fever, 0% vs 0% (NS); pulmonary events, 0% vs 0% (NS), for VNB+CDDP and VDS+CDDP arms, respectively

continued

Comments

Methodological comments

- Randomisation was centralised, and treatment arms were allocated using a computer-generated list stratified by institute, histology (squamous and non-squamous) and PS (0–1 and 2). An interim analysis after 16 months showed that the difference in response between the two groups had reached a level of significance sufficient for termination of patients' enrolment ($p = 0.004$, $\chi^2 = 8.378$), and the trial closed after accrual of 210 patients
- All prognostic factors were equally distributed between the two treatment arms
- Blinding was not specified
- To assess differences between proportions, p values were calculated by the χ^2 distribution (two-tailed tests) or Wilcoxon test. The Brook–Meyer–Crowley method was used to calculate the CIs for response. A proportional hazards regression analysis was used to determine the effects of different variables on response. Six variables (age, gender, histology, PS, stage and treatment arms) were analysed. A score was assigned to each variable for regression analysis. Survival curves were computed according to the Kaplan–Meier method, and differences in survival were compared with the log-rank and Wilcoxon tests for censored data. For interim analysis, O'Brien/Fleming multiple testing was employed to make up a three-step evaluation (40 patients/group) after the completion of evaluation for tumour response. Based on the interim analysis, the Effectiveness and Safety Committee discussed continuation of the trial
- Sample size calculations were based on an average response rate of 15% observed with VDS as a single agent and 25% observed with VNB as a single agent. Approximately 120 patients per arm were needed to detect a difference of 10–25% in response rates with a probability of 80% and a type I error of 5%
- In total, 210 patients were accrued for randomisation; however, 3 patients (1.4%) were ineligible because they had received previous chemotherapy. Of the remaining 207 patients, 1 patient withdrew consent and 2 patients had protocol violations. There were 204 patients assessable for response, toxicity and survival

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: measurable lesions, as shown by conventional chest X-ray or CT scan of the chest; no peripheral neuropathy; no severe constipation or ileus; adequate blood cell counts (Hb ≥ 9.5 g/dl, WBC ≥ 4000/μl and platelets ≥ 100,000/μl); adequate renal and liver functions (serum creatinine, ALT, AST and alkaline phosphatase ≤ twice the upper limit of the normal); and no evidence of severe heart or pulmonary diseases
- Interventions were clearly defined. VNB only, but if patients failed to respond after 4 cycles of monotherapy, they were switched to combination therapy of VDS+CDDP; VDS only, but if patients failed to respond after 4 cycles of monotherapy, they were switched to combination therapy of VNB+CDDP. Outcome measures were clearly defined
- Initial evaluation included complete history and physical examination, bronchofibroscopic examination, bone scintigraphy, abdominal CT scan or echography, brain CT, routine blood chemistries, blood cell counts, urinalysis and ECG. Every week, patients underwent physical examination, routine blood chemistries and complete blood cell counts. Follow-up study of the roentgenographic examination was performed with posteroanterior chest X-ray on all patients every week, but if lesions were not measurable on the posteroanterior chest X-ray, conventional tomography or CT scan was used
- *Conflicts of interest*: Supported by Kyowa Hakka Company

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	3/210 patients (1.4%) accrued for randomisation were ineligible because they had received previous chemotherapy 1/207 patients (0.5%) withdrew consent 2/207 patients (1%) had protocol violations

Reference and design	Intervention	Participants	Outcome measures
Le Chevalier *et al.*, 1994[43] (International) Phase III, international, multicentre, randomised trial Funding: Pierre Fabre Jadad score: 3/5	***VNB+CDDP arm*** VNB 30 mg/m^2 (iv) weekly and CDDP 120 mg/m^2 (iv) on days 1 and 29, then every 6 weeks ***VDS+CDDP arm*** VDS 3 mg/m^2 (iv) weekly for 6 weeks, then every 2 weeks, plus CDDP 120 mg/m^2 (iv) on days 1 and 29, then every 6 weeks ***VNB arm*** VNB 30 mg/m^2 (iv) weekly ***Length of treatment*** At least 10 weeks, unless progression was documented after 4 weeks. Patients with stable disease continued treatment for at least 18 weeks. Patients achieving an objective response continued treatment until toxicity or progression ***Other interventions used*** Saline, mannitol, antiemetics	*n* = 612 patients VNB+CDDP arm: *n* = 206 VDS+CDDP arm: *n* = 200 VNB arm: *n* = 206 90% of patients were male 79% of patients had PS of 0 or 1 Predominant histological type was squamous cell carcinoma 12% of patients had received previous surgery and/or radiotherapy ***Characteristics of target population*** Inoperable Stage III or IV NSCLC (IUAC classification) Aged 75 years and younger Median age: VNB+CDDP arm, 59 years; VDS+CDDP arm, 59 years; VNB arm, 60 years PS of 0–2 (WHO) No prior chemotherapy No symptomatic brain metastases 1+ unirradiated measurable lesion See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival time: Interval from date of randomisation until date of death or last follow-up ***Secondary end-points*** Response to treatment: Complete response defined as complete disappearance of all objective disease; partial response defined as ≥ 50% decrease in all targeted measurable lesions. Complete response and partial response confirmed by second evaluation after at least 4 weeks. Response evaluated by WHO criteria at 10 and 18 weeks of treatment Duration of objective response: Calculated from start of treatment to documented disease progression Toxicity: WHO criteria, except neurotoxicity, graded by Gralla scale ***Length of follow-up*** Median follow-up duration of 26 months (range, 16–36 months) See *General comments* below

Results

- In total, 73 patients (12%) were still alive on 15 September 1992 (randomisation took place between June 1989 and May 1991)
- *Survival*: Survival curve shown; VNB+CDDP curve above the VDS+CDDP and VNB curves. VNB+CDDP arm had the highest estimated median survival duration of 40 weeks, with VDS+CDDP arm estimated at 32 weeks and VNB arm estimated at 31 weeks. Unadjusted log-rank test comparing survival of patients receiving VNB+CDDP vs VDS+CDDP, $p = 0.085$; VNB+CDDP vs VNB, $p = 0.045$. Centre adjusted log-rank survival comparison for VNB+CDDP vs VDS+CDDP, $p = 0.04$; VNB+CDDP vs VNB, $p = 0.01$. Proportional hazards regression model indicated that relative instant death risk of VNB+CDDP vs VDS+CDDP was 1.4 (95% CI, 1.1 to 1.8; $p = 0.006$), and relative instant death risk of VNB+CDDP vs VNB was 1.4 (95% CI, 1.1 to 1.7; $p = 0.02$)
- *Response*: In VNB+CDDP arm (192 of 206 assessable patients), 57 patients (30%) confirmed complete response or partial response; in VDS+CDDP arm (183 of 200 assessable patients), 35 patients (19%) confirmed complete response or partial response; in VNB arm (199 out of 206 assessable patients), 28 patients (14%) confirmed complete response or partial response. VNB+CDDP arm had a significantly higher response rate in comparison with VDS+CDDP arm (χ^2 test, $p = 0.02$) and VNB arm (χ^2 test, $p < 0.001$). Estimated difference in response rates for VNB+CDDP compared with VDS+CDDP was 11% (95% CI, 2% to 20%) and 16% for VNB+CDDP compared with VNB (95% CI, 8% to 24%). Median duration of objective response for VNB+CDDP arm was 9.2 months (range, 3.8–36.4 months), for VDS+CDDP arm was 9.9 months (range, 3.7–29.6 months) and for VNB arm was 7.8 months (range, 3.9–21.8 months). In logistic regression model, when applied to treatment comparisons (VNB+CDDP vs VDS+CDDP and VNB+CDDP vs VNB), treatment (VNB+CDDP) and PS (2 vs 0–1) were the only significant predictors of response rate
- *QoL*: No QoL measures were reported
- *Toxicity*: Main toxic events were neutropenia (78.7% vs 47.6% vs 53.2%), thrombocytopenia (2.9% vs 3.1% vs 0.0%), sepsis (4.3% vs 2.0% vs 3.4%), neurological events (7.0% vs 17.0% vs 9.0%), ototoxicity (2.4% vs 1.3% vs 0.0%), renal events (5.9% vs 4.0% vs 0.0%), hepatic events (1.5% vs 2.3% vs 0.5%), alopecia (31.9% vs 38.0% vs 14.4%), nausea/vomiting (58.0% vs 59.0% vs 12.4%), diarrhoea (11.1% vs 6.8% vs 4.0%), local reaction (8.7% vs 2.1% vs 10.0%), for the VNB+CDDP, VDS+CDDP and VNB arms, respectively. Myelosuppression (neutropenia) was the most frequent toxic event, observed after delivery of 21% of vinca alkaloids in the VNB+CDDP arm, with 8% in VDS+CDDP arm and 11% in VNB arm (required hospitalisation for documented sepsis in 20 cases, and death was treatment related in 2 patients from each treatment arm). VNB+CDDP caused a significantly higher rate of neutropenia than the other two regimens ($p < 0.001$). Grade III and IV neurotoxicity was twice as frequent in the VDS+CDDP arm in comparison with the other treatment arms ($p < 0.004$)

continued

Comments

Methodological comments

- *Allocation of treatment groups*: Patients were randomly allocated to three arms, and randomisation was centralised. Treatment arms were allocated using a computer-generated list stratified by centre and stage. Many centres did not include stage IIIA patients, therefore were unable to use initial stratification data for efficacy comparisons. No placebo control
- *Comparability of groups*: The baseline characteristics (prognostic factors) of the individuals were well balanced among the three arms
- *Blinding*: Panel of at least three experts (blinded to patient treatment assignment) verified eligibility criteria, staging and toxicity, and reviewed original X-rays to evaluate response in all cases
- *Data analysis*: The two CDDP arms were compared first, then the better of the two was compared with the VNB arm. All patients were included in the survival analysis in their allocated treatment group. Prestratification-adjusted χ^2 and log-rank tests were used to compare estimates of response rates and survival. Fisher's exact test was used to compare toxicities. Logistic regression was used to adjust for and determine the impact of prognostic factors and treatment on the response rate. Cox's proportional hazards regression model was used to adjust for and determine the impact of prognostic factors and treatment on survival. Backward elimination procedure was used in both regression analyses to select major predicting factors. All significance levels were two sided. Investigators, worried about the activity of VNB being lower then the CDDP-containing regimens, carried out an interim unplanned analysis, after inclusion of 323 patients (37% of whom had died). Four patients were lost to follow-up
- *Power of study*: Sample size calculation was based on an average median survival duration of 32 weeks observed with VDS+CDDP in advanced NSCLC. At least 190 assessable patients per arm were needed to detect a difference in median survival (exponential model of survival) of 32–44 weeks with a probability of 80% and type I error of 5%, assuming a mean follow-up time of 2 years
- *Attrition*: Overall, 4% of patients were considered ineligible. Of the 24 patients considered ineligible, 5 had cerebral metastasis, 2 had previous malignancy, 2 had errors in diagnosis, 5 had PS of 3, and 10 had no measurable lesion. Of these ineligible patients, 9 were from the VNB+CDDP arm, 11 from VDS+CDDP arm and 4 from the VNB arm. Four patients (VNB+CDDP arm, 3 patients; VDS+CDDP arm, 1 patient) were lost to follow-up

General comments

- Eligibility and inclusion criteria were clearly defined. Additional inclusion criteria included: no prior malignancy, except adequately controlled basal cell carcinoma of the skin; no symptomatic brain metastases; no pre-existing hearing loss; no uncontrolled infection; normal blood count, liver function and renal function
- Interventions were clearly defined. Clear definitions of outcome measures: survival time, toxicity and response evaluation. The article was unclear regarding when evaluation of outcome measures was undertaken during the study (stated that "other diagnostic procedures performed according to symptoms"; see following paragraph)
- All patients underwent the following investigations: complete clinical examination, FBC and chemistries, ECG, chest X-ray, fibre-optic bronchoscopy, liver ultrasound; "other diagnostic procedures" included brain, thoracic and abdominal CT scan, bone scan and audiogram
- Concern exists that the CIs for survival curve displayed in a figure were, however, not reported in the text
- *Conflicts of interest*: Study supported by Pierre Fabre

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	24 patients (4%) were considered ineligible: 9 (4%) from VNB+CDDP arm, 11 (6%) from VDS+CDDP arm and 4 (2%) from VNB arm Reasons: 5 patients had cerebral metastasis, 2 had previous malignancy, 2 had errors in diagnosis, 5 had PS of 3, and 10 had no measurable lesion

Reference and design	Intervention	Participants	Outcome measures
Lorusso *et al.*, 1995[44] (Italy) Phase III, multicentre, randomised trial Funding: Not specified Jadad score: 2/5	***VNB arm*** VNB 25 mg/m^2 (iv) weekly ***VNB+CDDP arm*** VNB 25 mg/m^2 (iv) on days 1 and 8, with CDDP 80 mg/m^2 (iv) on day 1 every 3–4 weeks ***Length of treatment*** Minimum of two courses (VNB+CDDP) of 8 weeks (VNB). If patient had complete response, partial response or stable disease, treatment continued; otherwise, patients in VNB arm were treated with a CDDP regimen and those in VNB+CDDP arm were treated with a second-line therapy ***Other interventions used*** Not specified	*n* = 69 patients VNB arm: *n* = 35 VNB+CDDP arm: *n* = 34 97% of patients were male 66% of patients had squamous cell carcinoma 66% of patients had stage III disease 53% of patients had PS of 0 or 1 ***Characteristics of target population*** Histological confirmation of inoperable NSCLC at diagnosis Aged between 18 and 75 years Median age: VNB arm, 62 years (range, 54–73 years); VNB+CDDP arm, 61 years (range, 46–72 years) No previous chemotherapy Life expectancy > 3 months PS ≤ 2 (ECOG criteria) If patient had been previously subjected to radiotherapy, complete recovery from all side-effects was necessary ***Setting*** Not specified	***Primary end-points*** Response: WHO criteria. Evaluation of response after 8 weeks of VNB therapy and 2 courses of VNB+CDDP therapy Toxicity: WHO criteria. Evaluation not specified (possibly before each therapy cycle) ***Secondary end-point*** Survival: Evaluation not defined

Results

- *Response*: No complete response achieved in either arm. Partial response, 4 patients (13%) vs 13 patients (42%); stable disease, 11 patients (35%) vs 8 patients (26%); progressive disease, 16 patients (53%) vs 10 patients (32%), for VNB and VNB+CDDP arms, respectively. Partial response rate for VNB+CDDP arm was significantly higher than for VNB arm (Fisher's exact test, $p = 0.038$)
- *Survival*: Survival curves shown; both curves intersect at approximately 18, 22, 30, 54, 66 and 74 weeks. Median survival time: VNB arm, 30 weeks; VNB+CDDP arm, 38 weeks (NS; no significance level provided)
- *Haematological toxicity*: Grades 3–4: leucopenia (23% vs 8%, $p = 0.26$), anaemia (0% vs 27%, $p < 0.01$), thrombocytopenia (4% vs 11%), neurotoxicity (4% vs 15%), renal toxicity (8% vs 15%), phlebitis (15% vs 11%), for VNB and VNB+CDDP arms, respectively

Comments

Methodological comments

- Randomisation of patients was stratified according to stage and institution
- No difference between the prognostic factors was evident between the treatment arms
- Blinding was not specified
- Only evaluable patients were included in the analysis
- Sample size/statistical power was not specified. Concern exists that the study was underpowered to detect significant differences in response criteria
- A total of 69 patients were initially recruited; however, 3 patients from VNB+CDDP arm and 4 from VNB arm were lost to follow-up. One patient suspended treatment due to phlebitis and was administered a combination of CDDP+VP-16

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: presence of at least one measurable lesion; WBC > 3800/µl and platelets > 110,000/µl; normal values of bilirubin, serum glutamate oxaloacetate transaminase, serum glutamate pyruvate transaminase and serum creatinine; absence of neuropathy or brain metastases
- Interventions were clearly defined. Outcome measures were not clearly defined
- Concern exists that only haematological toxicities were reported
- Clinical response was verified by clinical examinations, chest radiography, chest CT scan, hepatic ultrasonography or hepatic CT scan, performed before entering study and every 8 weeks for the VNB arm and after each two courses of VNB+CDDP
- *Conflicts of interest*: Not specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	69 patients were initially recruited 3/34 patients (9%) from VNB+CDDP arm and 4/35 patients (11%) from VNB arm were lost to follow-up 1/62 patients (2%) suspended treatment due to phlebitis and was administered a combination of CDDP+VP-16

Reference and design	Intervention	Participants	Outcome measures
Martoni *et al*., 1998[45] (Italy) Phase III, multicentre randomised trial Funding: Not specified Jadad score: 2/5	***EPI+CDDP arm*** EPI 120 mg/m^2 (iv) with CDDP 60 mg/m^2 on day 1 of every 21-day cycle ***VNB+CDDP arm*** VNB 25 mg/m^2 (iv bolus) on days 1 and 8, with CDDP 60 mg/m^2 on day 1, of every 21-day cycle ***Length of treatment*** Maximum of 12 cycles or a cumulative dose of EPI of 840 mg/m^2 ***Other interventions used*** Standard antiemetic treatment (including 5-HT$_5$ receptor antagonists)	*n* = 228 (212 eligible) EPI+CDDP arm: *n* = 112 (102 eligible) VNB+CDDP arm: *n* = 116 (110 eligible) 84% of patients were male 48% of patients had epidermoid carcinoma 36% of patients had adenocarcinoma 40% of patients had stage IIIB disease, and 42% had stage IV disease Median PS of 80 (range, 70–100; good/excellent) ***Characteristics of target population*** Histological or cytological diagnosis of unresectable measurable NSCLC Locally advanced or metastatic NSCLC Bi- or uni-dimensionally measurable lesions Aged ≤ 72 years Median age: EPI+CDDP arm, 62 years (range, 42–72 years); VNB+CDDP arm, 61 years (range, 42–72 years) PS ≥ 70 (Karnofsky criteria) No previous chemotherapy or radiotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** Objective response: WHO criteria. Timing of evaluation not specified ***Secondary end-points*** Toxicity: WHO criteria, evaluated before each therapy cycle QoL: Evaluated by monitoring PS and symptoms (i.e. pain, dyspnoea and cough) Time to progression: Calculation of time to progression not specified Survival: Calculation of survival not specified See *General comments* below ***Length of follow-up*** Not specified

Results

- *Objective response*: Complete response, 1% vs 2%; partial response, 31% vs 25%; stable disease, 41% vs 40%; progressive disease, 21% vs 29%, for EPI+CDDP vs VNB+CDDP arms, respectively. Overall response rate for EPI+CDDP arm was 32% (95% CI, 23% to 41%) vs 27% (95% CI, 18% to 34%) for VNB+CDDP; difference not statistically significant, $p = 0.45$. Response was not related to stage or histotype (no significance levels reported)
- *Median time to progression*: EPI+CDDP arm, 6 months; VNB+CDDP arm, 5 months. Median duration of remission: EPI+CDDP arm, 9 months (range, 4–22 months); VNB+CDDP arm, 8 months (range, 3–34+ months)
- *Survival*: Survival curves shown; EPI+CDDP curve was above VNB+CDDP curve at all points. Median survival time for EPI+CDDP arm was 10.5 months (95% CI, 9.4 to 11.5 months) and 9.6 months (95% CI, 8.4 to 10.8 months) for the VNB+CDDP arm. One-year survival: EPI+CDDP arm, 42%; VNB+CDDP arm, 39%. Two-year survival: EPI+CDDP arm, 15%; VNB+CDDP arm, 8%. No differences were statistically significant
- *QoL*: Karnofsky PS was evaluated before and after 3–4 cycles in 56 patients in EPI+CDDP arm and 54 patients in VNB+CDDP arm. A ≥ 10% improvement was recorded in 37% and 39% of patients, respectively. Symptoms (cough, pain, dyspnoea) were evaluated in 58 patients in EPI+CDDP arm and 55 patients in VNB+CDDP arm before treatment. Of these, 64% and 70% of patients in EPI+CDDP and VNB+CDDP arms, respectively, suffered from at least one symptom. An improvement in at least one symptom, without worsening or the appearance of another one, was recorded in 57% and 61% of symptomatic patients, respectively. The difference was not statistically significant
- *Treatment after chemotherapy*: Of 212 eligible patients, 15 patients (8 in EPI+CDDP arm and 7 in VNB+CDDP arm) were operated on after 4 cycles; represented by 4 stage IIIA patients, 8 stage IIIB patients and 3 stage IV patients (two M1 lung and one M1 cervical ipsilateral lymph nodes). The operations consisted of pneumonectomy in 6 patients and lobectomy in the other 9 patients. In a further 6 patients in EPI+CDDP arm, the operation consisted of only an exploratory thoracotomy. Fifteen patients with stage III disease had chest radiotherapy after chemotherapy (9 in EPI+CDDP arm and 6 in VNB+CDDP arm)
- Nine patients who did not undergo operations were still alive at the time of analysis in early 1998 (randomisation was between August 1992 and February 1996): 5 patients from EPI+CDDP arm (after 22–37 months) and 4 patients from VNB+CDDP arm (after 23–40 months). Out of 15 operated patients, 7 were still alive in early 1998: 4 patients in EPI+CDDP arm (after 49, 43, 41 and 28 months; all NED but last) and 3 patients in VNB+CDDP arm (after 35, 21 and 21 months; all NED)
- *Haematological toxicity*: Myelosuppression was the most frequent side-effect, but the intensity was moderate. Grades 3–4: leucopenia (38% vs 21%, $p = 0.01$), anaemia (8% vs 7%, NS), thrombocytopenia (6% vs 0%, $p = 0.02$), for EPI+CDDP and VNB+CDDP arms, respectively. Febrile neutropenia was observed in 5 patients (5%) in EPI+CDDP arm and 3 patients (3%) in VNB+CDDP arm
- *Non-haematological toxicity*: Grade 3 adverse events (only nausea/vomiting): EPI+CDDP arm, 8% compared with VNB+CDDP, 4% (NS). Grades 1–2: vomiting/nausea (52% vs 61%, NS), alopecia (88% vs 33%, $p = 0.001$), stomatitis (14% vs 9%, NS), hyperpyrexia (16% vs 10%), renal toxicity (27% vs 34%, NS), for EPI+CDDP and VNB+CDDP arms, respectively
- In 62 patients (EPI+CDDP arm, 40 patients; VNB+CDDP arm, 22 patients), evaluation of LVEF by MUGA scan was repeated after 4 cycles. A > 15% LVEF absolute decrease was observed in 9 (22.5%) and 3 (14%) of the patients in EPI+CDDP and VNB+CDDP arms, respectively (no statistically significant difference)

continued

Comments

Methodological comments

- Patients were randomised to the two groups after balancing by histotype (epidermoid, adenocarcinoma, others) and stage (IIIA, IIIB, IV, recurrent)
- The two groups were well balanced. No statistically significant difference was evident in distribution of age, gender, PS, histology or stage
- Blinding was not specified
- ITT analysis was used. Descriptive statistics were used to report the study population with respect to demography and baseline characteristics. Overall tumour response and incidence of adverse effects were analysed using the appropriate method of ordered or non-ordered categorical data. Tabulated data were compared by the Pearson's χ^2 test. CIs (95%) were calculated assuming a binomial distribution. Variables concerning duration of response and time-to-events (time to progression, survival) were analysed according to Kaplan–Meyer product-limit method
- Sample size was calculated on the basis of expected overall tumour response. An overall objective response of 25% for VNB+CDDP was assumed, and a difference of 20% in the complete response plus partial response rate between VNB+CDDP and EPI+CDDP was considered to be of clinical interest. Setting α of 0.05 and power $(1-\beta)$ of 0.8, one-tailed level of significance, a sample size of 94 evaluable patients per treatment arm was computed
- There were 16 ineligible patients excluded from the trial due to: LVEF < 50% or history of myocardial infarction or arrhythmias requiring permanent medication, liver metastases with > 3 cm diameter, second tumour, symptomatic brain metastases, PS = 60%. In addition, 14 eligible patients (7 in each arm) were not evaluable for objective response because: 7 patients refused treatment (6 patients in EPI+CDDP arm, including 2 who refused to start treatment and 4 who refused to continue after first cycle [3 patients] or second cycle [1 patient]; 1 patient in VNB+CDDP arm, who refused to continue after first cycle). One patient in VNB+CDDP arm violated protocol after first cycle. Six patients (4 in EPI+CDDP arm and 2 in VNB+CDDP arm) died early in first or second cycle; in 2 of these patients, death was considered to be toxicity related; in the other 4 cases, death occurred after first cycle (2 patients: due to pulmonary oedema and unknown cause) and after second cycle (2 patients: sudden non-toxicity-related death and unknown cause)

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: absence of symptomatic brain metastases; normal LVEF evaluated by radionuclide angiocardiography at rest or echography; adequate bone marrow reserve, and hepatic and renal functions; absence of liver metastases with diameter > 2 cm
- Interventions were clearly defined. Outcome measures were not defined clearly, with little or no information included regarding the timing of evaluations. QoL measures used in the trial were not standard
- All patients were staged according to standard protocol (X-rays and CT scan of chest, ultrasound or CT scan of upper abdomen, bronchoscopy, bone scan). Blood count, biochemical tests and ECG were carried out at entry and before each subsequent course. Interim blood counts were carried out once weekly during treatment. In patients with stage IIIA/B disease, a CT scan was repeated after 3–4 cycles in order to evaluate the possibility of a surgical resection
- *Conflicts of interest:* None

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	16 ineligible patients were excluded from the trial due to: LVEF < 50% or history of myocardial infarction or arrhythmias requiring permanent medication, liver metastases with > 3 cm diameter, second tumour, symptomatic brain metastases, PS = 60% 14 eligible patients (7 in each arm) were not evaluable for objective response: 7/14 patients refused treatment, including 6/7 in EPI+CDDP arm (2/6 refused to start treatment, and 4/6 refused to continue after first cycle [3 patients] or second cycle [1 patient]) and 1/7 in VNB+CDDP arm (refused to continue after first cycle), and 1/7 from VNB+CDDP arm violated protocol after first cycle 6/7 patients (4 in EPI+CDDP arm and 2 in VNB+CDDP arm) died early in first or second cycle; in 2 of these patients, death was considered to be toxicity related; in the other 4 cases, death occurred after first cycle (2 patients: due to pulmonary oedema and unknown cause) and after second cycle (2 patients: sudden non-toxicity-related death and unknown cause)

Reference and design	Intervention	Participants	Outcome measures
Pérol *et al*., 1996[46] (France) Phase II, open, multicentre, randomised trial Funding: Supported by Pierre Fabre Ltd Jadad score: 2/5	***Group A*** CDDP 120 mg/m^2 (iv) with MITO 8 mg/m^2 (iv) on days 1, 29 and 71, with VDS 3 mg/m^2/week (iv) for 5 weeks and then every 2 weeks up to the 15th week ***Group B*** CDDP 120 mg/m^2 (iv) with MITO 8 mg/m^2 (iv) on days 1, 29 and 71, with VNB 25 mg/m^2/week (iv) for 16 weeks ***Length of treatment*** Three courses of treatment for the full 16 weeks Mean duration of treatment was 11.7 weeks in Group A and 12.8 weeks in Group B ***Other interventions used*** Mannitol, sodium chloride, isotonic saline, dexamethasone, antiemetics, ondansetron (5-HT$_3$ receptor antagonist)	*n* = 227 patients Group A: *n* = 113 Group B: *n* = 114 95% of patients were male 77% of patients had PS of 0 or 1 63% of patients had squamous cell carcinoma 41% of patients had stage IIIB disease ***Characteristics of target population*** NSCLC confirmed by histological or cytological examination Stage III or IV inoperable disease No previous treatment Aged < 75 years Median age: Group A, 60 years; Group B, 61 years PS of 0–2 (WHO scale) At least one bi-dimensional measurable target Absence of cerebral metastases See *General comments* below ***Setting*** Not specified	***Primary end-points*** Response: Standard response criteria, evaluated after administration of three courses (i.e. 17th week) Survival: Measured from date of randomisation ***Secondary end-point*** Toxicity: WHO criteria (except neurotoxicity, for which Gralla scale was used) See *General comments* below ***Length of follow-up*** Median duration of follow-up was 39 months

Results

- *Response*: Response was evaluated only in those patients receiving full three courses of treatment. In Group A, 80 patients (71%) were evaluable, and in Group B, 83 patients (73%) were evaluable for response. Complete response, 2 patients (2%) vs 4 patients (4%); partial response, 17 patients (15%) vs 24 patients (21%); progressive disease, 33 patients (29%) vs 28 patients (25%); stable disease, 28 patients (25%) vs 27 patients (24%), for Group A and Group B, respectively (all NS). Overall response was not significantly different (p = 0.15) between Group A (17%; 95% CI, 10% to 24%) and Group B (25%; 95% CI, 17% to 32%). Response rate in stage III patients: Group A, 25% (95% CI, 15% to 36%); Group B, 28% (95% CI, 17% to 38%); p = 0.78. Response rate in stage IV patients: Group A, 6% (95% CI, 0% to 12%); Group B, 20% (95% CI, 8% to 31%); p < 0.05
- *Survival*: Overall survival curves shown; Group B curve above after about 10 weeks. Stage III survival curves shown; Group B curve above after 6 weeks. Stage IV survival curves shown; Group A and Group B curves comparable. Median survival time: overall, Group A, 33.4 weeks, and Group B, 34.5 weeks (log-rank test, p = 0.2); stage III , Group A, 33.4 weeks, and Group B, 45.9 weeks (log-rank test, p = 0.13); stage IV , Group A, 33.8 weeks, and Group B, 27.6 weeks (log-rank test, p = 0.902). No statistically significant difference was evident between two treatment groups. Overall 2-year survival: Group A, 15.6%, and Group B, 9%; p = 0.13. Stage III 1-year survival: Group A, 44.6%, and Group B, 26.2%; p = 0.03
- *Haematological toxicity*: Grades 3 and 4: anaemia (16% vs 31%, p < 0.01), thrombocytopenia (11% vs 7%, p = NS), neutropenia (61% vs 87%, p < 0.01), for Group A and Group B, respectively
- *Non-haematological toxicity*: Grades 3 and 4: renal events (3% vs 2%, p = NS), cardiac events (3% vs 2%, p = NS). Grades 2–4: sepsis (16% vs 29%, p < 0.03), neurological events (23% vs 6%, p < 0.01), nausea/vomiting (28% vs 21%, p = NS), local reaction (0% vs 7%, p < 0.05)

Comments

Methodological comments

- Patients were randomised to either Group A or B; no details of method used. Group A was the control arm. Patients were stratified by centre
- No significant difference between arms were evident with respect to age, gender, histology type, stage or PS
- Open study. The initial staging, inclusion criteria and responses were systematically evaluated by a panel of experts who were unaware of the treatment received
- Not ITT analysis for response (only patients who had received the three courses of treatment were evaluated for treatment), but ITT for survival and toxicity. The chi-squared test was employed to compare the rates of response and the toxicities. Survival curves were constructed using actuarial methods. Survivals were compared using log-rank test. The statistical procedures were carried out using Epi Infor and PCSM software. Significance levels were two-sided. An intermediate analysis was carried out for toxicity after inclusion of the first 100 patients
- Main objective of trial was to compare survival and objective response rates. Size of population was based on an expected increase in response rate from 30% to 50%. With 111 patients per arm, the chance of demonstrating this difference with a 5% chance of a type I error was 80%
- In total, 240 patients were recruited, with 119 in Group A and 121 in Group B, but 13 patients were excluded as not meeting inclusion criteria. Three were 33 patients (29%) in Group A and 31 patients in Group B who did not receive the full three courses of treatment, mostly due to toxicity progression. These patients were treated as treatment failures in the evaluation of response. Group A: 24/33 patients discontinued due to treatment toxicity, 5/33 patients refused treatment, 1/33 were lost to follow-up, and 3/33 had intercurrent illness. Group B: 21/31 patients discontinued due to treatment toxicity, 1/31 patients refused treatment, 2/31 were lost to follow-up, 6/31 had intercurrent illness, and 1/31 violated the protocol. The difference between groups was not significant. Four patients in Group A and 2 patients in Group B were lost to follow-up

continued

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: satisfactory renal and hepatic function (serum creatinine < 140 µmol/l with clearance > 60 ml/minute if creatinine > 110 µmol/l; serum bilirubin < 35 µmol/l); no superior vena cava syndrome; no history of neoplastic disease, except localised uterine cervical cancer or basal cell carcinoma of skin; no infection; peripheral neuropathy ≥ 2 on Gralla scale; absence of cardiac or symptomatic coronary insufficiency
- Interventions were clearly defined. Outcome measures were clearly defined. A complete response was defined as the disappearance of all measurable or evaluable lesions, with negative signs in endobronchial samples for a period of at least 4 weeks. A partial response corresponded to a > 50% reduction in the sum of the areas of the measurable lesions, without progression to other sites, in the absence of the appearance of other tumoural sites for 4 weeks
- Pretreatment evaluation: full clinical examination, laboratory investigations including blood count, electrolytes, hepatic enzymes, creatinine clearance if serum creatinine > 110 µmol/l, front and side chest X-rays, ECG, fibre-optic bronchoscopy, chest CT scan with sections on adrenal zones, CT scan or ultrasonographic examination of abdomen, brain CT scan, bone scintigraphy, biopsy of hepatic or adrenal lesion, if this was sole metastatic site contraindicating surgical resection of the primary tumour. Response was evaluated after administration of the courses on the 17th week
- *Conflicts of interest*: Study supported by Pierre Fabre Oncologie (Boulogne, France)

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	33 patients (29%) in Group A and 31 patients in Group B did not receive the full three courses of treatment Group A: 24/33 patients (73%) discontinued due to treatment toxicity, 5/33 patients (15%) refused treatment, 1/33 patients (3%) were lost to follow-up, and 3/33 patients (9%) had intercurrent illness Group B: 21/31 patients (68%)discontinued due to treatment toxicity, 1/31 patients (3%) refused treatment, 2/31 patients (6%) were lost to follow-up, 6/31 patients (19%) had intercurrent illness, and 1/31 patients (3%) violated the protocol Four patients in Group A and 2 patients in Group B were lost to follow-up

Reference and design	Intervention	Participants	Outcome measures
Wozniak *et al.*, 1998[47] (USA – SWOG) Phase III, multicentre randomised trial Funding: Glaxo Wellcome Jadad score: 2/5	***VNB+CDDP arm*** CDDP 100 mg/m^2 every 4 weeks and VNB 25 mg/m^2 weekly ***CDDP arm*** CDDP 100 mg/m^2 every 4 weeks ***Length of treatment*** Not stated ***Other interventions used*** Pretreatment iv hydration, mannitol diuresis and post-treatment hydration, poly-antiemetic regimens	*n* = 415 eligible patients VNB+CDDP arm: *n* = 206 CDDP arm: *n* = 209 67% of patients were male 92% of patients had stage IV disease 52% of patients had adenocarcinoma ***Characteristics of target population*** Histological/cytological stage IIIB or IV NSCLC Median age: VNB+CDDP arm, 63 years (range, 33–84 years); CDDP arm, 63 years (range, 37–81 years) Stage IIIB patients ineligible for combined modality therapy (chemotherapy plus radiotherapy) PS of 0 or 1 (scale not indicated) No brain metastases Bi-dimensionally measurable or assessable disease No previous chemotherapy Two weeks must have elapsed since major surgery and 3 weeks since radiotherapy ***Setting*** Inpatient and outpatient	***Primary end-point*** Survival: Standard SWOG criteria used, although evaluation time not defined (12- and 24-month survival rates reported) ***Secondary end-points*** Response: Standard SWOG criteria used. SWOG policy is to perform a second response confirmation 4 weeks after initial response is determined Toxicity: Standard SWOG criteria used (evaluation time not stated) ***Length of follow-up*** Not stated, but 2-year survival data reported

Results

- No information was provided on survival at end of study period (randomisation between November 1993 and April 1995). One patient remained on study in September 1996
- *Survival*: Survival curves shown; VNB+CDDP arm above CDDP arm. Median progression-free survival: 4 months for VNB+CDDP arm and 2 months for CDDP arm (p = 0.0001). Median overall survival: 8 months for VNB+CDDP arm vs 6 months for CDDP arm (p = 0.0018). Twelve-month survival: 36% for VNB+CDDP arm vs 20% for CDDP arm; 24-month survival: 12% for VNB+CDDP vs 6% for CDDP. Survival analysis indicated that stage IV disease (median survival of 13 months for stage IIIB vs 7 months for stage IV; p = 0.003) and elevated lactate dehydrogenase (p = 0.0001) had a negative impact on survival
- *Response*: Complete response, 2% vs 0%; partial response, 24% vs 12%; stable disease, 47% vs 34%; progressive disease, 27% vs 54%, for VNB+CDDP arm and CDDP arm, respectively. Overall response was 26% (8% unconfirmed) for VNB+CDDP arm vs 12% (4% unconfirmed) for CDDP arm (p = 0.0002)
- *QoL*: No QoL measures were reported
- *Haematological toxicity*: Main toxicity was myelosuppression. Grades 3 and 4: granulocytopenia (22% and 59% vs 4% and 2%), thrombocytopenia (4% and 2% vs 2% and 1%), anaemia (21% and 3% vs 8% % 0.5%), for VNB+CDDP and CDDP arms, respectively
- *Non-haematological toxicity*: Grades 3 and 4 combined: nausea/vomiting (20% vs 20%), malaise/weakness (15% vs 11%), electrolyte imbalance (6% vs 3%), renal events (5% vs 5%) and deaths (5 patients vs 1 patient), for VNB+CDDP and CDDP arms, respectively. The 5 deaths in the VNB+CDDP arm were attributable to 2 cases of cardiac ischaemia, 1 massive cerebrovascular accident, 1 multisystem failure as a result of a VNB overdose, and 1 due to catheter sepsis and progressive disease
- Dose-delivered intensity: Median courses: three with VNB+CDDP vs two with CDDP

Comments

Methodological comments

- Patients were randomised to one of two study arms. Patients were stratified (at time of registration) by lactate dehydrogenase level (normal vs abnormal), disease stage, disease status (measurable vs unmeasurable), prior surgical resection or radiotherapy, and histology. Stratification was available by stage for survival outcome only. No placebo control
- Baseline characteristics were well balanced between both arms
- No indication of blinding
- ITT was not used for response rates. All patients were included in the survival analysis in their allocated treatment group. CIs were not given. All patients were included in 2-year survival calculations
- A sample size of 400 patients was needed to have 88% power to detect the stated alternative with a 5% significance level. VNB+CDDP was considered superior if it demonstrated a 33% increase in survival and 50% increase in 1-year survival. Survival curves were estimated by the product-limit method and compared using the log-rank test, stratified by predetermined prognostic factors. Cox regression was used to explore the influence of prognostic factors on survival and to assess treatment-by-factor interactions
- Seventeen patients (4%) were ineligible: 10 due to inappropriate stage; others due to inadequate or inappropriate pathology, brain metastases, inadequate disease assessment or administration of radiotherapy within 3 weeks of registration

continued

General comments

- *Deviations from protocol*: 10 deviations in CDDP arm (7 received no protocol treatment, 1 received 1 day of protocol treatment, and 2 received radiotherapy) and 18 deviations in VNB+CDDP arm (2 received no protocol treatment, 11 received 1 day of protocol treatment, 2 were treated when treatment should have been withheld, and 1 received 5 times the dose of VNB)
- Eligibility and inclusion criteria were clearly defined; however, amendment of protocol part-way through allowed stage IIIB patients into trial. Additional inclusion criteria included: neutrophil count ≥ 1500/μl, platelet count ≥ institutional limits of normal, Hb ≥ 9 g/dl, serum creatinine concentration ≤ 1.5 mg/dl or calculated clearance ≥ 60 ml/minute, bilirubin level ≤ 2.0 mg/dl, and AST ≤ twice the institutional upper limits of normal unless liver metastases were present, in which case ≤ 4 times the institutional upper limits of normal permitted; concomitant radiotherapy was not allowed
- Interventions were clearly defined. Definitions of outcome measures were not clear; also no information was provided on time of evaluation for response, survival or toxicity. No information was provided on evaluation of general health status prior to, or during, the trial
- Survival curves were displayed as figures, CIs were not reported
- *Conflicts of interest*: Supported by grant from Glaxo Wellcome

Quality assessment for RCTs (Jadad score[19])

Question	**Score**
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	17 patients (4%) were ineligible after registration (no data given regarding to which arms they were randomised) 10 patients had inappropriate stage; others had either inadequate or inappropriate pathology, brain metastases, inadequate disease assessment or administration of radiotherapy within 3 weeks of registration

Reference and design	Intervention	Participants	Outcome measures
ELVIS Group, 1999[48] (Italy) Phase III, multicentre, randomised trial Funding: Not specified Jadad score: 3/3	***VNB arm*** VNB (iv) 30 mg/m^2 on days 1 and 8 of 21-day cycle ***BSC arm*** Participating investigators free to choose the treatment strategy ***Length of treatment*** Maximum of 6 cycles ***Other interventions used*** Palliative radiotherapy in both arms, second-line treatment or crossover not allowed, antiemetic treatment (metoclopramide 20 mg) given as iv bolus before VNB	*n* = 154 patients VNB arm: *n* = 76 patients BSC arm: *n* = 78 patients 87% of patients were male 73% of patients had stage IV NSCLC 76% of patients had ECOG PS of 0–1 46% of patients had squamous histology type ***Characteristics of target population*** Histologically or cytologically confirmed NSCLC Stage IIIB or IV (with pleural effusion or metastatic supraclavicular lymph nodes) NSCLC Aged ≥ 70 years Median age: VNB arm, 74 years (range, 70–85 years); BSC arm, 74 years (range, 70–86 years) PS ≤ 2 (ECOG scale) No previous chemotherapy See *General comments* below ***Setting*** Not specified	***Primary end-point*** QoL: EORTC QLQ-C30 and LC13 evaluated at baseline, before each therapy cycle (VNB) and before each follow-up visit (BSC) ***Secondary end-points*** Response (VNB only): Standardised criteria evaluated before each therapy cycle (response not confirmed after ≥ 4 weeks) Toxicity (VNB only): WHO criteria evaluated before each therapy cycle Survival: Defined as the interval from the date of randomisation to the date of death or last follow-up See *General comments* below ***Length of follow-up*** Potential follow-up of at least 18 weeks. Median follow-up for 28 alive patients was 57 weeks (range, 10–106 weeks)

Results

- *QoL*: Results of QoL analysis were determined by mixed effect model. In total, 141 patients (91.6%) completed baseline QoL questionnaires. There was no difference in baseline scores of the QoL scales between the two treatment arms in all of the analysed domains. Suboptimal values for global health status were found in 36.8% (VNB arm) and 31.5% (BSC arm) of patients. Sixty patients (39%) completed questionnaires after 6 cycles. EORTC functional scales were consistently better for the patients receiving VNB, although statistical significance was reached only for cognitive function (p = 0.02) and was borderline for global health status (p = 0.06). Among symptom scales, VNB patients scored clearly better than control patients for one item (pain, p = 0.02), but scored significantly worse for some treatment toxicity-related items (constipation, p = 0.002; nausea and vomiting, p = 0.07). For the LC13, VNB patients performed better for dyspnoea (p = 0.05), pain medication (p = 0.01) and pain in shoulder (p = 0.12), but worse than the BSC arm for peripheral neuropathy (p = 0.04) and hair loss (p = 0.0001)
- *Survival*: As of 30 June 1998, 126 (82%) of patients had died (59 in VNB arm and 67 in BSC arm; randomisation was between April 1996 to June 1997). Survival curves shown. The VNB survival curve was above the control curve at all points on curve (log-rank test, p = 0.03; Cox model, p = 0.02). Median survival: VNB arm, 28 weeks; BSC arm, 21 weeks. Six-month survival rate: VNB arm, 55%; BSC arm, 41%. Twelve-month survival rate: VNB arm, 32%; BSC arm, 14%. After adjustment by stage of disease and ECOG PS, the estimated relative hazard of death for patients receiving VNB compared with the control arm was 0.65 (95% CI, 0.45 to 0.93)
- *Response*: VNB arm: complete response, 1/76 patients (1%); partial response, 14/76 patients (18%); stable disease, 23/76 patients (30.3%); progressive disease, 32/76 (42.1%). Objective response rate: 19.7% (95% CI, 11.5% to 30.5)
- *Toxicity*: There were 71 patients assessed for toxicity. Treatment was stopped in 5 patients due to severe toxic events: WHO grade 3 constipation in 3 patients, grade 4 constipation in 1 patient and grade 2 heart toxicity in 1 patient. Other relevant toxic events, not including treatment stopping, were grade 4 leucopenia (1), grade 4 neutropenia (3), grade 3 vomiting (1) and grade 3 alopecia (3)
- *Discontinuations*: Treatment was discontinued in 11 patients (2 after one course, 1 after two courses, 4 after three courses, 3 after four courses, 1 after five courses). In total, 283 courses of chemotherapy were delivered (median, 4 cycles per patient). One patient received a reduced dose (VNB 20 mg)

Comments

Methodological comments

- Patients were randomly assigned to treatment centrally at the Data Coordinating Centre, using a computer-driven minimisation procedure, and using the centre, stage of disease at entry and the ECOG PS as stratification variables
- Patients were well matched for pretreatment characteristics: age, gender, ECOG PS and stage
- ELVIS investigators, who halted the trial, were blind to results. Otherwise blinding was unspecified
- Effect of VNB on QoL was evaluated by fitting a linear mixed model for each EORTC scale. An unstructured (fully parameterised) within-subject covariance matrix was used in analysis. BMDP-5V software was used for QoL analysis. All p-values were two sided. Survival curves were drawn up using the Kaplan–Meier product limit method and compared by using the Mantel-Haenszel test. The relative hazard of death and 95% CI were estimated by the Cox proportional hazard model using treatment, stage of disease and ECOG PS as covariates
- QoL was the primary end-point. Therefore, the study had power of 80% to recognise an effect of 30%; at third assessment after patients were randomly assigned, approximately 350 patients were needed
- In control arm, 3/78 patients received chemotherapy. In VNB arm, data on chemotherapy data were missing for 3/76 patients; 2 patients never received chemotherapy because they died within 1 week of being randomly assigned. Chemotherapy was discontinued earlier than planned in 11 patients (2 after one course, 1 after two courses, 4 after three courses, 3 after four courses and 1 after five courses), and treatment was stopped in 5 patients due to toxicity. In the 30 of the 55 patients who fully respected the protocol, chemotherapy was stopped before the sixth course due to progressive disease

continued

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: no presence of brain metastases, no previous history of another cancer (excluding non-melanomatous skin cancer and *in situ* cervical cancer), no reduced bone marrow or renal or hepatic function
- Interventions were clearly defined. Outcomes were clearly defined
- *Response definitions*: Complete response, disappearance of all known sites of disease; partial response, decrease of 50% or more in the sum of the products of largest perpendicular diameters of measurable lesions, with no appearance of new lesions and no progression of lesions; stable disease, decrease of < 50% or increase of > 25% in sum of products of largest perpendicular diameter of measurable lesions and no progression of any lesion; progressive disease, ≥ 25% increase in the size of 1+ measurable lesions or appearance of a new lesion
- Prior to study entry, clinical examination included a two-view chest X-ray, CT scan of the thorax and abdomen, and a bone scan for assessment of disease extension. Before each dose, patients in VNB arm underwent a clinical examination that included a routine biochemistry work-up and blood cell counts; patients in control arm were scheduled for six clinical evaluations, one evaluation every 3–4 weeks. Response was evaluated after third and sixth cycles. QoL was assessed at baseline (before treatment), and before each course of chemotherapy in VNB arm and before each follow-up visit in control arm
- Concern exists for lack of safety data reported
- *Conflicts of interest*: None specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	Not possible due to comparison with BSC
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	3/78 patients in control arm received chemotherapy. In VNB arm, data on chemotherapy data was missing for 3/76 patients; 2 patients never received chemotherapy because they died within 1 week of being randomly assigned. Chemotherapy was discontinued earlier than planned in 11 patients (2 after one course, 1 after two courses, 4 after three courses, 3 after four courses and 1 after five courses); treatment was stopped in 5 patients due to toxicity. In the 30 of the 55 patients who fully respected the protocol, chemotherapy was stopped before the sixth course due to progressive disease

TNM, pathological staging scheme for tumours (T, primary tumour; N, regional nodes; M, metastatic); BUN, blood urea nitrogen; IUAC, International Union Against Cancer; M1, metastatic 1 (subclassification of pathological staging); NED, not defined in reference but presumed to be no evidence of disease; MUGA, not defined in reference but presumed to be multiple gated acquisition (blood pool)

Appendix 9

Summary of evidence of effectiveness of combined therapies in lung cancer

Reference and design	Intervention	Participants	Outcome measures
Comella *et al*., 1999[49] (Italy) Phase II, randomised trial Funding: Not specified Jadad score: 2/5	***Group A (CDDP+GEM+ VNB arm)*** CDDP 50 mg/m^2 (iv), GEM 1000 mg/m^2 (iv), with VNB 25 mg/m^2 (iv) on days 1 and 8 of every 3-week cycle ***Group B (CDDP+EPI+VDS+ LON arm)*** CDDP 80 mg/m^2 (iv), EPI 80 mg/m^2 and VDS 3 mg/m^2 (iv) all delivered on day 1 every 4 weeks, with LON (oral) 150 mg/m^2 three times daily Dose modified or delayed according to *a priori* criteria depending on toxicity ***Length of treatment*** All patients had 3 cycles of treatment; patients with complete response or partial response received an additional 3 cycles ***Other interventions used*** Normal saline, anti-emetic prophylaxis (hydroxytryptamine-3 receptor antagonists plus dexamethasone) Radio-therapy after 3 cycles for patients with stable disease and 6 cycles for grade III responders	*n* = 111 patients Group A: *n* = 57 Group B: *n* = 54 90% of patients were male 56% of patients had stage IV disease 60% of patients had squamous histology 77% had PS of 70–80 ***Characteristics of target population*** Histologically or cytologically confirmed, locally advanced (stage IIIB) or metastatic NSCLC Aged ≤ 70 years Median age: Group A, 65 years (range, 40–70 years); Group B, 64 years (range, 38–70 years) No previous chemotherapy or thoracic radiotherapy PS ≤ 1 (WHO scale) or PS ≥ 70 (Karnofsky scale) Life expectancy ≥ 12 weeks See *General comments* below ***Setting*** Not specified	***Primary end-point*** Response: WHO criteria, evaluated after 3 therapy cycles ***Secondary end-points*** Toxicity: WHO scale, evaluated before each therapy cycle Survival: Measured from date of entry to date of death or last known follow-up evaluation QoL: Not defined as an end-point but assessed in patients at two centres. Ten-item questionnaire (derived from LCSS), administered at diagnosis, 3 and 6 cycles, and thereafter every 3 months until death See *General comments* below ***Length of follow-up*** Median follow-up duration of 82 weeks (range, 52–121 weeks)

continued

Results

- *Response*: Complete response, 4/87 patients (5%), 1/57 (2%), 2/54 (4%); partial response, 46/87 patients (53%), 33/57 (58%), 18/54 (33%); stable disease, 16/87 patients (18%), 12/57 (21%), 14/54 (26%); progressive disease, 21/87 patients (24%), 11/57 (19%), 20/54 (37%), for whole Group A, randomised Group A and Group B, respectively. Overall response: 50/87 patients (57%; 95% CI, 46% to 68%), 34/57 (60%; 95% CI, 46% to 72%), 20/54 (37%; 95% CI, 24% to 51%), for entire Group A, randomised Group A and Group B, respectively. Response according to stage: stage IIIB disease was associated with a higher response rate in both arms (63% and 41% in Groups A and B, respectively), compared with stage IV disease (52% and 34%, in Groups A and B, respectively; NS). Response according to PS and histology: a slightly higher, though not statistically significantly different, response rate was observed in Group A for patients with an ECOG PS of 0 (67% vs 54% for ECOG PS of 1) and squamous histology (63% vs 50% for those with other histologies), whereas overall response rate did not vary according to PS and histology in the control arm. Improvement in PS: Group A, 48/57 patients (55%), Group B, 20/54 (37%)
- *Survival*: Survival curves shown; Group A curve is higher than Group B curve, with stage IIIB curves being higher than Stage IV, and response curve being higher than non-response. After median follow-up of 82 weeks, 31 patients were still alive and 25 were progression-free in Group A. Overall survival duration: Group A, 50 weeks (95% CI, 41 to 58 weeks); Group B, 33 weeks (95% CI, 24 to 41 weeks). One-year survival probabilities: Group A, 48%; Group B, 29%. Two-year survival probabilities: Group A, 19%; Group B, 0%. Median progression-free survival: Group A, 32 weeks (95% CI, 27 to 38 weeks); Group B, 18 weeks (95% CI, 11 to 26 weeks). Median survival was longer in stage IIIB patients than in those with stage IV disease (61 vs 47 weeks for Group A and 42 vs 29 weeks for Group B), and in those with PS of 0 compared with 1 (62 vs 46 weeks for Group A and 36 vs 22 weeks for Group B). Response to treatment was identified with longer survival duration in both arms (62 vs 39 weeks for Group A and 41 vs 26 weeks for Group B). Median survival of non-responders in Group A was similar to that of responders in Group B. Cox analysis performed on the whole population failed to identify any pretreatment features that significantly affected survival. When response to treatment was included in the Cox model as a time-dependent covariate after stratification for treatment, it was strongly predictive of a longer survival outcome. The relative risk of dying was 1.98 (range, 1.26 to 3.12) in non-responders compared with responders ($p = 0.004$)
- *QoL*: QoL was evaluated in 74/141 patients (Group A, 51; Group B, 23). The QoL score improved in 30/51 patients (59%) and 9/23 patients (39%) in Groups A and B, respectively. Among the 39 patients with improved QoL scores, 35 had an objective response and 37 showed symptom relief
- *Haematological toxicity*: Grades 3–4: neutropenia (46% vs 22%), thrombocytopenia (14% vs 11%), anaemia (10% vs 13%), for Group A and Group B, respectively
- *Non-haematological toxicity*: Grades 3-4: vomiting (5% vs 11%), renal events (0% vs 0%), neuropathy (0% vs 0%), mucositis (0% vs 0%), alopecia (7% vs 13%), fatigue (6% vs 6%), diarrhoea (0% vs 0%), arthralgias (0% vs 2%), abdominal pain (0% vs 2%), for Group A and Group B, respectively
- No toxic deaths occurred. Only 5 patients (3 in Group A and 2 in Group B) discontinued treatment early because of severe toxicity. Nine patients in Group A were hospitalised due to neutropenic fever (10%), compared with 2 patients (4%) in Group B. Overall, 3 patients (3%) in Group A and 1 (2%) in Group B required platelet transfusions due to symptomatic thrombocytopenia. Severe anaemia requiring PRBC transfusions was similar in both arms (almost 10% of patients)

Comments

Methodological comments

- Randomisation occurred centrally; method was not stated. Randomisation was stopped early in December 1996. Not placebo controlled. However, Group B was the control arm, and Group A was the experimental arm. Stratification was by stage (IIIB vs IV) and centre. Overall, 87 patients were enrolled into Group A and 54 into Group B. Unequal distribution occurred because randomisation was stopped in December 1996 (57 and 54 patients had been enrolled, respectively). At this point, the minimum number of responses required to consider the experimental treatment worth further evaluation in a Phase III trial had been reached, but it was decided to continue experimental treatment to better estimate activity. Group B accrual stopped because it was considered unethical to treat patients with a regimen that had been shown to be remarkably lower in activity
- Although last 30 patients in Group A were not randomised, the two arms were fairly well balanced for the main pretreatment characteristics. A slightly lower prevalence of stage IIIB disease occurred in Group B, in which more patients showed a squamous histotype compared with Group A. However, these differences were not statistically significant. Demographic characteristics appeared similar, although not statistically tested
- Blinding was not specified
- Patients not assessable for response were considered as non-responding in the calculation of objective response rate; 95% CIs were calculated for response rates. Patients receiving at least 1 cycle of chemotherapy were considered eligible for QoL evaluation. Analysis was conducted on the basis of randomisation (ITT) and on the follow-up data available as of 28 February 1998. Survival curves were estimated using the Kaplan–Meier product-limit method. A Cox model was used for multivariate analysis of assigned treatment, as well as of the effect on survival of pretreatment variables, including gender (male vs female), age (< 60 years vs others), PS (0 vs 1), stage (IIIB vs IV) and histotype (squamous vs others) as covariates
- Initially, a target overall response rate (p_1) for the experimental combination of 45% and a lowest response rate of interest (p_0) of 30% were chosen. Simon two-stage optimal design was used; a total of 81 patients were needed to test the hypothesis, but at least 10 major responses had to be observed in the first 27 assessable patients to continue enrolment up to the final sample size (with type 1 and 2 errors of 0.05 and 0.20). However, to avoid the treatment of too high a number of patients with a regimen of little therapeutic interest, another analysis was planned after the first 46 assessable patients, because at least 17 objective responses were needed to be observed to accept the hypothesis
- Between October 1995 and March 1997, a total of 145 patients were entered into the study, of which 141 were eligible (63 patients with stage IIIB disease and 78 patients with stage IV disease). Two patients (in each arm) were excluded because of major protocol violations (3 patients with PS of 2, and 1 patient with Hb level < 100 g/l). Only 5 patients (3 in Group A and 2 in Group B) discontinued treatment early because of severe toxicity. It was not stated whether exclusions were from the 57 patients in Group A

continued

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate bone marrow function (absolute neutrophil count ≥ 2 x 10^9/l, platelet count ≥ 100 x 10^9/l, Hb level ≥ 100 g/l); adequate liver function (bilirubin level < 2 times the upper limit of the normal, AST and /or ALT < 3 times the upper limit of normal, prothrombin time < 1.5 times control) and creatinine clearance ≥ 60 ml/minute; no CNS metastases; no severe cardiac arrhythmia or heart failure; no second- or third-degree heart block, or acute myocardial infarction within 4 months before study entry
- Interventions were clearly defined. Outcome measures were clearly defined. Response: complete response was defined as the complete disappearance of disease at all sites; partial response was defined as a reduction of at least 50% in the sum of the products of the longest diameters of all measurable lesions, with no appearance of new lesions; PS and symptom assessments (ECOG and Karnofsky scale); QoL was limited to non-validated questionnaire on a subset of patients; for survival, the date of randomisation was considered the date of entry; progression-free survival was defined as time elapsed from the date of entry to the date of progressive disease or death without progression
- Pretreatment evaluation included complete history and physical examination, ECG, chest radiograph, respiratory test, fibre-optic bronchoscopy, and CT scan of chest and upper abdomen. Radionuclide scan of bone and CT scan of brain were also performed as necessary to document disease extent. Laboratory investigation included a complete blood cell count with WBC differential and platelet count, a full chemistry profile, prothrombin time, partial thromboplastin and thrombin time, and urinalysis. For each chemotherapy course, patients underwent physical examination, laboratory tests and chest radiograph; all diagnostic procedures required to evaluate response to treatment were performed after the third and sixth course; weekly blood cell count was performed; complete tumour response assessments were performed after 3 cycles, with a minimum of 4 weeks required to document response and the best response recorded for each patient. Toxicity was assessed before each cycle of therapy, and haematological assessments were carried out weekly to determine toxicity at the nadir. For toxicity analysis, the worst data for each patient in all cycles of chemotherapy were used. PS (ECOG or Karnofsky) and symptom assessments were performed before each cycle of therapy. QoL was not an end-point but was assessed in two participating centres, using a 10-item questionnaire, derived from the LCSS, at diagnosis, 3 and 6 cycles, and thereafter every 3 months until death
- *Conflicts of interest:* Not specified

Quality assessment for RCTs (Jadad score[19])

Question	**Score**
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	Between October 1995 and March 1997, a total of 145 patients were entered into the study, of which 141 were eligible 2 patients in each arm were excluded because of major protocol violations (3 patients with PS of 2, and 1 patient with Hb level < 100 g/l) 3/87 patients (3%) in Group A and 2/54 patients (4%) in Group B discontinued treatment early because of severe toxicity

Reference and design	Intervention	Participants	Outcome measures
Comella *et al*., 2000[50] (Italy) Phase III, randomised trial (interim analysis) Funding: Not specified Jadad score: 3/5	***Group A*** CDDP 50 mg/m^2 (iv), GEM 1000 mg/m^2 (iv), with VNB 25 mg/m^2 (iv) on days 1 and 8 every 3 weeks ***Group B*** CDDP 100 mg/m^2 (iv) on day 1, with GEM 1000 mg/m^2 (iv) on days 1, 8 and 15 every 4 weeks ***Group C*** CDDP 120 mg/m^2 (iv) on days 1 and 29 (then every 6 weeks), with VNB 30 mg/m^2/week for 10 weeks ***Length of treatment*** Not specified ***Other interventions used*** Antiemetic prophylaxis (5-hydroxytryptamine-3 receptor antagonists plus dexamethasone)	*n* = 180 patients Group A: *n* = 60 Group B: *n* = 60 Group C: *n* = 60 93% of patients were male 72% of patients had PS of 1 49% of patients had squamous histology 58% of patients had stage IV disease ***Characteristics of target population*** Histologically or cytologically confirmed locally advanced (stage IIIB) or metastatic (stage IV) NSCLC Aged ≤ 70 years Median age: Group A, 62 years (range, 38–70 years); Group B, 61 years (range, 35–70 years); Group C, 60 years (range, 38–70 years) No previous chemotherapy No prior surgery or thoracic radiotherapy PS ≤ 1 (ECOG criteria) Life expectancy ≥ 12 weeks See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Measured from the date of entry (day of randomisation) until the date of death or date of last follow-up ***Secondary end-points*** Response: WHO criteria, evaluated after a similar time from initial therapy had elapsed for each patient (after third course for Group A and the second course for Groups B and C) Toxicity: WHO criteria, assessed before each therapy cycle QoL: 10-item questionnaire derived from the LCSS, evaluated before each therapy cycle ***Length of follow-up*** Minimum potential follow-up of 24 weeks. Last follow-up data available as of 15 April 1999. Median potential follow-up of 16 months (range, 6–24 months) See *General comments* below

Results

- *Survival*: As of 15 April 1999 (randomisation started in April 1997), 128 patients had died (33, 42 and 53 patients in Groups A, B and C, respectively). Median survival time: Group A, 51 weeks; Group B, 42 weeks; Group C, 35 weeks. One-year survival probabilities: Group A, 45%; Group B, 40%; Group C, 34%. Median survival time (by disease stage): 48 weeks for stage IIIB and 40 weeks for stage IV ($p > 0.05$). Median survival time (by age): 47 weeks for age < 65 years and 41 weeks for age > 65 years ($p > 0.05$). Stage IV median survival time: Group A, 47 weeks; Group B, 34 weeks; Group C, 27 weeks. Cox analysis on survival: including age, ECOG PS, histology and weight loss as covariates, the estimated hazard of death for patients in Group A compared with Group B was 0.35 (95% CI, 0.16 to 0.77)
- *Response*: Complete response, 3% vs 0% vs 0%; partial response, 43% vs 30% vs 25%; stable disease, 28% vs 37% vs 25%; progressive disease, 25% vs 33% vs 50%, for Groups A, B and C, respectively. Overall response rate: Group A, 47% (95% CI, 34% to 60%); Group B, 30% (95% CI, 19% to 43%); Group C, 25% (95% CI, 15% to 38%). Stratified overall response rates (stratified for stage IIIB vs IV): Group A, 54% vs 42%; Group B, 36% vs 26%; Group C, 29% vs 22%
- *Response with second-line therapy*: In total, 95 patients received second-line chemotherapy (34 vs 31 vs 30 patients, in Groups A, B and C, respectively); 15 objective responses (16%) were recorded with second-line chemotherapy, without substantial differences in overall response rate between patients coming from each of the three arms
- *Haematological toxicity*: Grades 3–4: neutropenia, 45% vs 40% vs 75% ($p < 0.001$); thrombocytopenia, 17% vs 30% vs 20%; anaemia, 15% vs 13% vs 25%
- *Non-haematological toxicity*: Grades 3–4: vomiting, 15% vs 30% vs 50% ($p < 0.0001$); mucositis, 2% vs 5% vs 1%; diarrhoea, 1% vs 3% vs 1%; fatigue, 14% vs 10% vs 15%, for Groups A, B and C, respectively
- *Chemotherapy discontinued early due to toxicity*: Three patients in Group A discontinued therapy early due to lack of haematological recovery (1), severe emesis (1), refusal due to severe fatigue (1); 10 patients in Group B due to toxic deaths (2), severe renal toxicity (1), severe constipation (2), severe emesis (3), severe fatigue (2); 3 patients in Group C due to lack of recovery from renal toxicity (1), lack of haematological recovery (1), severe emesis (1)

continued

Comments

Methodological comments

- Randomisation took place from April 1997, to one of three treatment arms. Patients were analysed in April 1999. Randomisation was performed centrally, by computer-driven minimisation procedure using centre and stage (IIIB vs IV) at entry as stratifying variables
- Baseline characteristics were well balanced across the three arms
- Blinding was not specified
- Cox model was used to produce the *p*-values because a treatment comparison was provided adjusting for the effects of the main prognostic factors. In this model, all the pretreatment characteristics that were expected to affect survival were included as covariates, which were age (< 65 years vs older), PS (0 vs 1), stage (IIIB vs IV), histology (squamous vs others) and weight loss (> 5% vs < 5% of the body weight). Analysis was conducted on the basis of randomisation (ITT analysis) and on the follow-up data available as of 15 April 1999. Survival curves were estimated using the Kaplan–Meier product-limit method
- Survival was the primary end-point of study. The study targeted to have a power of 80% to recognise a 50% prolongation of the median survival time in Group A when compared with Groups B and C (12 months vs 8 months). Selecting an alpha error of 0.05 (two-tailed test), approximately 120 patients for each arm were needed. According to the two-stage design for Phase III trials, an interim analysis was planned when the first 60 patients per arm were assessable for survival. The study had to be discontinued if Group A did not show a reduction in the risk of death reaching a *p*-value < 0.44 when compared with either Group B or C (null hypothesis accepted). Early discontinuation of accrual was also planned in any doublet regimen showing a significantly increased risk of death ($p < 0.01$) when compared with Group A (null hypothesis rejected)
- On 16 January 1999, 218 patient had been enrolled. Six patients did not meet eligibility criteria due to malignancy other than NSCLC (2), incorrect staging (2), absence of adequate PS or bone marrow reserve (2); five patients were not assessable because of complete absence of any demographic and follow-up information. A total of 207 patients (Group A, 69; Group B, 70; Group C, 68) were eligible and assessable; an interim analysis was carried out on the first 60 patients in each arm. Five patients (8%) in Group A were not reassessed because of early progression, death or treatment discontinuation by any cause, compared with 8 patients (13%) in Group B and 20 patients (33%) in Group C

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate bone marrow function (absolute neutrophil count ≥ 2 x 10^9/l, platelet count ≥ 100 x 10^9/l, Hb level ≥ 100 g/l); adequate liver function (bilirubin level 2 times the upper limit of normal, AST and/or ALT < 3 times the upper limit of normal, prothrombin time < 1.5 times control) and creatinine clearance ≥ 60 ml/minute; ECOG PS ≤ 1; no presence of severe cardiac arrhythmia or heart failure; no second- or third-degree heart block, or acute myocardial infarction within 4 months before study entry; no CNS metastases
- Interventions were clearly defined. Outcome measures were clearly defined
- Pretreatment evaluation included complete history and physical examination, ECG, chest X-ray, respiratory tests, fibre-optic bronchoscopy, and CT scan of chest, brain and upper abdomen. Radionuclide scan of bone was also performed to document disease extent. Laboratory investigations included a complete blood cell count with the WBC differential and platelet count, a full chemistry profile, prothrombin time, partial thromboplastin and thrombin time, and urinalysis. For each chemotherapy course, physical examination, laboratory tests and chest X-ray were performed. Weekly complete blood cell count was performed. All the diagnostic procedures required to evaluate response to treatment were scheduled after a similar time from initial therapy (after the third course in Group A and after the second course in Groups B and C). Tumour response assessment: all patients underwent complete tumour response assessment after a similar time from initial therapy (after 3 cycles in Group A and after 2 cycles in Groups B and C). A minimum duration of 4 weeks was required to document a response. Toxicity was assessed before each cycle of chemotherapy, and haematological assessments were performed weekly to determine the toxicity at the nadir. For toxicity analysis, the worst data for each patient in all cycles of chemotherapy were used. PS and symptom assessments were performed before each cycle of chemotherapy. Complete response was defined as the disappearance of disease at all sites, and partial response was defined as a reduction of at least 50% in the sum of the products of the longest diameters of all measurable lesions, with no appearance of new lesions
- *Conflicts of interest:* None specified

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	6/218 patients did not meet eligibility criteria due to malignancy other than NSCLC (2), incorrect staging (2), absence of adequate PS or bone marrow reserve (2) 5/218 patients were not assessable because of complete absence of any demographic and follow-up information 5 patients (8%) in Group A were not reassessed because of early progression, death or treatment discontinuation by any cause, compared with 8 patients (13%) in Group B and 20 patients (33%) in Group C 3 patients in Group A discontinued therapy early due to lack of haematological recovery (1), severe emesis (1), refusal due to severe fatigue (1); 10 patients in Group B due to toxic deaths (2), severe renal toxicity (1), severe constipation (2), severe emesis (3), severe fatigue (2); 3 patients in Group C due to lack of recovery from renal toxicity (1), lack of haematological recovery (1), severe emesis (1)

Reference and design	Intervention	Participants	Outcome measures
Kosmidis *et al*., 2000[51] (Greece) Phase III, randomised trial (preliminary results) Funding: Not specified Jadad score: 2/5	***Group A*** PAX 200 mg/m^2 (iv), with CBDCA (an area under concentration–time curve of 6 according to the Calvert formula; iv on day 1; PAX given before CBDCA; repeated every 3 weeks ***Group B*** PAX 200 mg/m^2 (iv), with GEM 1000 mg/m^2 (iv) on days 1 and 8; PAX given before GEM; repeated every 3 weeks ***Length of treatment*** Minimum of 2 cycles and a maximum of 6 cycles ***Other interventions used*** Anti-hypersensitivity prophylaxis, antiemetic therapy (ondansetron)	*n* = 127 patients Group A: *n* = 63 Group B: *n* = 64 91% of patients were male 51% of patients had PS of 0 91% of patients were smokers 91% of patients had no prior surgery 83% of patients had no prior radiotherapy ***Characteristics of target population*** Histologically documented, inoperable, recurrent or metastatic NSCLC (inoperable stage IIIA, IIIB or IV carcinoma; AJCC criteria) Aged ≥ 18 years Median age: Group A, 63 years (range, 45–78 years); Group B, 65 years (range, 38–78 years) Life expectancy ≥ 12 weeks PS of 0 to 2 (ECOG criteria) Measurable or evaluable disease in non-radiated fields, unless subsequent disease was documented See *General comments* below ***Setting*** Not stated	***Primary end-point*** Response rates: WHO criteria. Time of evaluation not provided ***Secondary end-points*** Time to progression: Calculation of time to progression not provided Survival: Calculation of survival time not provided Toxicity: WHO criteria. Time of evaluation not provided QoL: EORTC-QLQ-C30. Time of evaluation not provided See *General comments* below ***Length of follow-up*** Median follow-up time of 4.6 months

Results

- *Response*: Complete response, 0% (95% CI, 0% to 7.7%) vs 4.2% (95% CI, 0.5% to 14.3%); partial response, 21.8% (95% CI, 9.4% to 33.9%) vs 33.3% (95% CI, 20.4% to 48.4%); stable disease, 47.8% (95% CI, 34.9% to 65.1%) vs 37.5% (95% CI, 23.9% to 52.7%); progressive disease, 30.4% (95% CI, 17.1% to 45.8%) vs 25% (95% CI, 13.6% to 39.6%), for Group A and Group B, respectively. Difference between groups was not statistically significant (Kruskal-Wallis exact test, $p = 0.12$; Fisher's exact test, $p = 0.07$)
- *Time to progression*: Not reported
- *Survival*: Not reported
- *QoL*: Not reported
- *Haematological toxicity*: Grades 3–4: anaemia (1.6% vs 1.6%), neutropenia (9.6% vs 3.1%), thrombocytopenia (1.6% vs 0%), for Group A and Group B, respectively. No statistical difference was found
- *Non-haematological toxicity*: Grades 3–4: nausea/vomiting (0% vs 0%), diarrhoea (1.6% vs 0%), alopecia (30.2% vs 28.2%), neurotoxicity (4.8% vs 0%), myalgia/arthralgia (0% vs 1.6%), neutropenic fever (0% vs 0%), cutaneous reaction (0% vs 1.6%), fatigue (0% vs 0%), arrhythmia (0% vs 1.6%), for Groups A and B, respectively. It is worth noting that grade 4 alopecia, diarrhoea and neurotoxicity occurred in 1 patient, each in Group A; only 6 patients in Group B suffered from grade 4 alopecia

Comments

Methodological comments

- Patients were randomised to either Group A or Group B; method not stated
- Characteristics of the patients, tumour characteristics and sites of metastases were comparable
- Blinding was not specified
- Method of data analysis was not specified
- Sample size/statistical power was not specified
- In Group A, 17 (26.9%) of patients were not evaluable, for the following reasons: too early (10), voluntary withdrawal (5) and death (heart failure) (1); note that 1 patient is missing. In Group B, 16 (25%) of patients were not evaluable, for the following reasons: too early (11), voluntary withdrawal (3), death (heart failure) (1), discontinued (1)

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: function tests must be within the following ranges, WBC ≥ 3500/µl, platelet count ≥ 100,000/µl, creatinine ≤ 1.4 mg/dl, creatinine clearance > 60 ml/minute, bilirubin ≥ 1.2 mg/dl, and normal alkaline and gamma-glutamyl transferase; no active infections; no history of other neoplasms, except for basal carcinoma of skin or *in situ* carcinoma of the cervix; not previously treated with cytotoxic chemotherapy; no active cardiac disease or pre-existing motor or sensory neuropathy, WHO grade 3 or 4
- Interventions were clearly defined. Outcome measures were not defined clearly
- Disease staging and response were assessed by clinical examination, chest X-ray and CT scans. Abdominal CT scans, liver or adrenal ultrasound tests, and bone scans were used as indicated. Disease parameters were measured at least every 8 weeks; chest X-rays were repeated monthly. All histopathology slides were reviewed by the same panel of pathologists. No other details were provided
- Preliminary results did not include information on time to progression, survival or QoL
- *Conflicts of interest*: Lead author has received honoraria from Eli Lilly and Company

continued

Quality assessment for RCTs (Jadad score[19])

Question	**Score**
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	Group A, 17 (26.9%) of patients were not evaluable, for the following reasons: too early (10), voluntary withdrawal (5), death (heart failure) (1); note that 1 patient is missing Group B, 16 (25%) of patients were not evaluable, for the following reasons: too early (11), voluntary withdrawal (3), death (heart failure) (1), discontinued (1)

Reference and design	Intervention	Participants	Outcome measures
Perry *et al.*, 2000[52] (USA) Phase II, randomised trial Funding: In part, grants from NCI to CALGB Jadad score: 2/5	***PAX+IFOS arm*** PAX 250 mg/m^2 (iv) on day 1, with IFOS (iv) 1.6 g/m^2 on days 1, 2 and 3, repeated every 21 days ***VNB+IFOS arm*** VNB 30 mg/m^2 (iv) on days 1, 2 and 3, with IFOS (iv) 1.6 g/m^2 on days 1, 2 and 3, repeated every 21 days ***Length of treatment*** Maximum of 6 cycles ***Other interventions used*** Mesna (iv) with IFOS, G-CSF	*n* = 93 patients PAX+IFOS arm: *n* = 48 VNB+IFOS arm: *n* = 45 63% of patients were male 84% of patients had stage IV or recurrent disease 32% of patients had PS of 0 67% of patients had a weight loss of < 5% ***Characteristics of target population*** Histological documentation of NSCLC, stage IIIB or IV disease Aged > 18 years Median age: 63 years (range, 32–81 years) No prior chemotherapy Measurable or evaluable disease PS of 0–1 (CALGB criteria) See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Definition of how survival time was calculated not provided ***Secondary end-points*** Failure-free survival: Definition of how failure-free survival was calculated not provided Response: 'Standard' criteria (not specified). Time of evaluation not provided Toxicity: Criteria not specified. Time of evaluation not provided ***Length of follow-up*** Not specified

Results

- *Response*: Complete response, 2 patients (4%) vs 2 patients (4%); partial response, 16 patients (33%) vs 11 patients (24%), for PAX+IFOS and VNB+IFOS arms, respectively. Overall response: PAX+IFOS arm, 18 patients (38%; 95% CI, 24% to 53%); VNB+IFOS arm, 14 patients (31%; 95% CI, 18% to 47%). One patient in VNB+IFOS arm had a regression of evaluable disease. No significance levels were provided
- *Survival*: Survival curve shown; VNB+IFOS curve intersected PAX+IFOS curve several times and was roughly comparable. Median survival time: PAX+IFOS arm, 8.5 months; VNB+IFOS arm, 7.4 months (95% CI, 5.3 to 13.3 months). One-year estimated survival: PAX+IFOS arm, 35% (95% CI, 24% to 52%); VNB+IFOS arm, 38% (95% CI, 26% to 55%)
- *Median failure-free survival*: PAX+IFOS arm, 4.8 months; VNB+IFOS arm, 2.4 months
- *Haematological toxicity*: Grades 3–4: WBC (48% vs 98%), granulocytes/bands (59% vs 93%), lymphocytes (67% vs 73%), Hb (13% vs 20%), hyperglycaemia (10% vs 0%), for PAX+IFOS and VNB+IFOS arms, respectively
- *Non-haematological toxicity*: Infection (17% vs 24%), dyspnoea (14% vs 4%), sensorimotor events (8% vs 13%), constipation (2% vs 11%), pain (2% vs 15%), malaise/fatigue (12% vs 22%), for PAX+IFOS and VNB+IFOS arms, respectively
- Treatment-related failures (death): PAX+IFOS arm, 1 due to respiratory failure; VNB+IFOS arm, 1 due to neutropenic fever

Comments

Methodological comments

- Patients were randomised to either PAX+IFOS arm or VNB+IFOS arm; method not stated
- Very limited description of patient characteristics, with no comparison between groups
- Blinding was not specified
- Kaplan–Meier curves were used to describe survival and failure-free survival. Toxicities were tabulated by the most severe occurrence experienced by each individual patient
- For each treatment regimen, the study was designed to differentiate between a 14% level of activity ($p < 0.15$) and a 30% level of activity ($p \geq 0.3$). Forty-five patients were to be accrued to each arm, and if fewer than 10 patients responded (complete response, partial response or regression of evaluable disease), the null hypothesis would be accepted, and it would be concluded that the arm did not have sufficient activity to merit further investigation. The study was designed with $\alpha = \beta = 0.1$, that is, the probability of concluding that the treatment is ineffective ($p \leq 0.14$) when in reality $p \geq 0.3$ (or vice versa) is 0.10
- In total, 100 patients were accrued. Two patients were ineligible because PS was 2, and 5 patients never received protocol treatment (withdrew consent [1], ineligible after study registration but before initial treatment [3] or unknown reasons [1]). Two treatment-related deaths occurred: PAX+IFOS arm, 1 (respiratory failure); VNB+IFOS arm, 1 (neutropenic fever)

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: granulocytes > 1800/μl, platelets > 100,000/μl, creatinine < 2 times the upper limit of normal, bilirubin < 1.5 times the upper limit of normal; no CNS metastases; not, if stage IIIB disease, eligible for other CALGB protocols of combined chemotherapy or radiotherapy; no history of prior cancer; no other serious medical/psychiatric illnesses
- Interventions were clearly defined. Outcome measures were not defined clearly. Response: complete response required disappearance of all measurable or evaluable disease, signs, symptoms and biochemical changes related to the tumour, lasting at least 4 weeks post-therapy, during which time no new lesions could appear; partial response required a reduction of ≥ 50% in the sum of the products of the perpendicular diameters of all measurable lesions, lasting at least 4 weeks post-therapy, during which no new lesions should appear; regression of evaluable disease implied a definite decrease in tumour size agreed upon by two independent investigators, and no new lesions for > 8 weeks. Toxicity: most severe occurrence experienced by each individual patient
- Concern exists for the lack of information relating to patient characteristics. Significance levels have not been provided
- *Time of assessment*: Not specified
- *Conflicts of interest*: Supported in part by grants from the NCI to the CALGB

continued

Quality assessment for RCTs (Jadad score[19])

Question	**Score**
Was the study described as randomised?	1
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	1
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	100 patients were accrued; 2 patients were ineligible because PS = 2; 5 patients never received protocol treatment (withdrew consent [1], ineligible after study registration but before initial treatment [3], unknown reasons [1]) 2 treatment-related deaths: PAX+IFOS arm, 1 (respiratory failure); VNB+IFOS arm, 1 (neutropenic fever)

Reference and design	Intervention	Participants	Outcome measures
Frasci *et al.*, 2000[53] (Italy) Phase III, multicentre, randomised trial Funding: No funding information provided Jadad score: 2/5	***GEM+VNB arm*** GEM 1200 mg/m^2 (iv) and VNB 30 mg/m^2 on days 1 and 8 of 3-week cycle ***VNB arm*** VNB 30 mg/m^2 on days 1 and 8 of 3-week cycle ***Length of treatment*** Maximum of 6 cycles. Chemotherapy suspended after 3 cycles in the absence of a major objective response ***Other interventions used*** All patients received antiemetic prophylaxis that consisted of HT_3 receptor antagonists	*n* = 120 GEM+VNB arm: *n* = 60 VNB arm: *n* = 60 90% of patients were male 24% of patients had PS of 2 48% of patients had squamous carcinoma 39% of patients had adenocarcinoma histology 59% of patients had stage IV NSCLC ***Characteristics of target population*** Aged ≥ 70 years Median age: GEM+VNB arm, 75 years (range, 71–83 years); VNB arm, 74 years (range, 71–81 years) Histologically or cytologically confirmed, locally advanced or metastatic NSCLC (stage IIIB and IV) No prior chemotherapy or thoracic radiotherapy PS ≤ 2 (ECOG scale) Life expectancy ≥ 12 weeks See *General comments* below ***Setting*** Not specified	***Primary end-point*** Survival: Calculated from the date of randomisation until the date of death or last follow-up ***Secondary end-points*** Response rate: WHO criteria, evaluated after 3 chemotherapy cycles (additional 3 cycles for patients with complete or partial response), then every 2 months until death Toxicity: WHO criteria, evaluated before each therapy cycle QoL: Modified LCSS questionnaire. Completed at diagnosis, after 3 and 6 cycles, and thereafter every 2 months until death See *General comments* below ***Length of follow-up*** Minimum of 12 weeks

Results

- *Survival*: Survival curves shown; GEM+VNB curve above VNB curve from approximately week 10. As of 31 May 1999, 93 patients had died (GEM+VNB arm, 41; VNB arm, 52). Median survival time was 29 weeks for GEM+VNB arm and 18 weeks for VNB arm. The projected 6-month and 12-month survival rates were 56% and 30% in the GEM+VNB arm compared with 32% and 13% in the VNB arm, respectively (Mantel-Haenszel test, $p < 0.01$,). In multivariate Cox analysis, after adjusting for disease stage, PS, histology, Charlson score and weight loss, the risk of death in GEM+VNB arm compared with VNB arm was 0.48 (95% CI, 0.29 to 0.79)
- *Response*: No complete responses were observed; partial response, 22% vs 15%; stable disease, 27% vs 12%; progressive disease, 22% vs 25%, for GEM+VNB and VNB arms, respectively. Overall response rate: GEM+VNB arm, 22% (95% CI, 12% to 34%); VNB arm, 15% (95% CI, 7% to 27%). In stage IIIB and IV disease, overall response rates were 25% and 19% in the GEM+VNB arm and 20% and 11% in the VNB arm, respectively. A total of 18 patients (30%) in the GEM+VNB arm and 29 (48%) in the VNB arm were not restaged because of early progression or death or treatment discontinuation by any cause
- *QoL*: In total, 111 (92%) of patients completed baseline questionnaires (no differences in baseline scores between the two arms); 81% (34 of 43 alive) completed the assessment at 6 months. Almost 60% of patients in the GEM+VNB arm did not show impairment of the QoL score during the treatment, compared with approximately 40% in the VNB arm. Fourteen (26%) of patients in the GEM+VNB arm showed symptom relief during the treatment, compared with 8 (15%) in the VNB arm. In particular, cough (31% vs 17%) and shortness of breath (28% vs 11%) were more frequently improved in the combination group. The probability of being alive at 6 months without symptom deterioration was 43% vs 22% for the GEM+VNB and VNB arms, respectively
- *Haematological toxicity*: Grades 3 and 4: neutropenia (30% and 8% vs 23% and 5%), thrombocytopenia (10% and 3% vs 7% and 2%), anaemia (7% and 0% vs 2% and 0%), for GEM+VNB arm and VNB arm, respectively
- *Non-haematological toxicity*: Grades 3 and 4: vomiting (13% and 2% vs 5% and 0%), alopecia (7% and 0% vs 0% and 0%), for GEM+VNB arm and VNB arm, respectively

continued

Comments

Methodological comments

- Patients were randomised centrally to the two treatment arms by a computer-driven minimisation procedure that used the centre, stage (IIIB vs IV) and PS (ECOG 0 or 1 vs 2) as stratifying variables
- Baseline characteristics of patients were comparable in all three arms, for age, gender, PS, extent of involvement and histology
- Blinding was not specified
- ITT analysis was used for survival. Survival curves were estimated by the Kaplan–Meier product-limit method. Multivariate analysis of the effect of the different effects of treatment on survival was performed by a Cox model and included PS, stage, histotype, Charlson score and weight loss as covariates. For QoL, the time to symptom deterioration was calculated in each arm by using the Kaplan–Meier product-limit method
- The study planned to have a power of 80% to recognise a 50% prolongation of the median survival time in the GEM+VNB arm compared with the VNB arm (9 vs 6 months, respectively). An interim analysis was planned after 60 patients per arm had been enrolled, with a minimum potential follow-up of 12 weeks, according to Schaid's two-stage design for survival comparisons. The study would have been stopped if the GEM+VNB arm had failed to show a reduction in the risk of death with a p-value of at least 0.44 (null hypothesis accepted). The study would have also been stopped if patients who received VNB alone had shown an increased risk of death, with $p > 0.01$, when compared with the GEM+VNB arm
- As of 31 May 1999, 152 patients were enrolled. In total, 125 eligible patients (GEM+VNB arm, 63; VNB arm, 62) were included in the study of 23 February 1999, these patients having the 12-week minimum potential for follow-up at the time of analysis. Five patients were excluded (GEM+VNB arm, 3; VNB arm, 2) because no information had been sent to the coordinating centre. Therefore, 120 patients were eligible and assessable

General comments

- Eligibility and exclusion criteria were clearly described. Additional inclusion criteria included: adequate bone marrow and liver function; no severe cardiac arrhythmia or heart failure, second- or third-degree heart block, and acute myocardial infarction within 4 months of study entry
- Interventions and outcome measures were clearly described
- Pretreatment evaluation included a complete history and physical examination, ECG, chest X-ray, respiratory tests, fibre-optic bronchoscopy, and CT scans of the chest, brain and upper abdomen. Radionuclide scans of bone were also performed as necessary to document disease extent. An accurate evaluation of co-morbidities was also used before starting chemotherapy, and a score was calculated for each patient by using the Charlson scale. Physical examination, laboratory tests and chest X-ray were performed at each chemotherapy cycle, and all diagnostic procedures to evaluate response to treatment were performed in both arms after the third and sixth cycles and every 2 months until death. Complete blood cell count was performed weekly. For toxicity, PS and symptom assessments were performed before each cycle of chemotherapy
- The article was based on an interim analysis. The study was not powered to detect significant differences between treatment groups in terms of survival. Concern exists that the reporting of QoL was insufficient
- *Conflicts of interest:* None stated

Quality assessment for RCTs (Jadad score[19])

Question	Score
Was the study described as randomised?	2
Was the study described as double-blind?	0
Was there a description of withdrawals and dropouts?	0
What proportion of sample (intervention and control groups separately) withdrew or dropped out?	No information provided

CALGB, Cancer and Leukemia Group B; G-CSF, granulocyte colony-stimulating factor

Appendix 10

Characteristics of gemcitabine, vinorelbine, paclitaxel and docetaxel economic evaluation studies

	GEM	GEM	GEM	GEM
Study number	1	2	3	4
Authors	Copley-Merriman *et al.*[56]	Copley-Merriman *et al.*[57]	Evans[55]	Evans[61]
Publication year	1996	1996	1997	1996
Base year prices	1990–91, 1992, 1993	?	1993	1993
Intervention	GEM: 1000–1250 mg/m^2 CDDP+VP-16: CDDP 60 mg/m^2 or 100 mg/m^2, and VP-16 100 mg/m^2	Different perspectives used for three models ***US model*** GEM: 1250 mg/m^2 CDDP+VP-16: CDDP 60 mg/m^2 or 100 mg/m^2, and VP-16 100 mg/m^2 ***Spanish model*** GEM: 1250 mg/m^2 CDDP+VP-16: CDDP 60 mg/m^2 or 100 mg/m^2 or 25–50 mg/m^2, and VP-16 100 mg/m^2 or 120 mg/m^2 ***German model*** GEM: 1250 mg/m^2 IFOS+VP-16: IFOS 1500 mg/m^2 (with mesna 20%) and VP-16 120 mg/m^2	GEM: starting from 1000 mg/m^2 weekly x 3 followed by a week's rest BSC	GEM: starting from 1000 mg/m^2 weekly x 3 followed by a week's rest BSC
Study type	Cost-minimisation model (Monte Carlo simulation)	Cost minimisation (Monte Carlo)	CEA model – population health model (microsimulation model)	CEA model – population health model (microsimulation model)
NSCLC stage	III and IV	IIIA, IIIB and IV	IV	IV
Perspective	Private insurance payer	Private insurance payer, government as payer	Government as payer in universal healthcare system	Government as payer in universal healthcare system
Industry role	Authors work for producer of drug	Authors work for producer of drug	Drug producer is funder, and author also works for them	Funder of study
Country of origin	USA	USA, Spain, Germany	Canada	Canada

continued

	GEM and VNB	GEM	GEM	GEM
Study number	5	6	7	8
Authors	Palmer & Brandt[59]	Koch *et al.*[88]	Sacristan *et al.*[60]	Tennvall & Fernberg[58]
Publication year	1996	1995	2000	1998
Base year prices	?	?	1997	1996
Intervention	GEM+CDDP, VNB+ CDDP, MITO+IFOS+ CDDP, VP-16+CDDP Different regimens used because this study used a meta-analysis	GEM: 1250 mg/m^2 IFOS+VP-16: IFOS 1500 mg/m^2 (with mesna 20%) and VP-16 120 mg/m^2	GEM+CDDP: GEM 1250 mg/m^2 (iv) on days 1 and 8, plus CDDP 100 mg/m^2 on day 1 of each 21-day cycle, before GEM infusion VP-16+CDDP: VP-16 100 mg/m^2 on days 1, 2 and 3, plus CDDP 100 mg/m^2 on day 1 of each 21-day cycle, before VP-16 infusion	GEM: 1000 mg/m^2 VP-16+CDDP: VP-16 100 mg/m^2 and CDDP 120 mg/m^2 IFOS+VP-16: IFOS 1500 mg/m^2 (with mesna 20% 300 mg/m^2) and VP-16 120 mg/m^2
Study type	CEA model – influence model and decision tree	Cost minimisation	Cost minimisation/ CEA	Cost minimisation
NSCLC stage	IIIB and IV	IIIA, IIIB and IV	IIIB and IV	IIIB and IV
Perspective	Third-party payer	Direct hospital costs	Government as payer	Societal
Industry role	None ?	Co-authors work for pharmaceutical company	Funder and principal author from pharma-ceutical company	Funder – acknowledgement
Country of origin	Italy	Germany	Spain	Sweden

continued

	VNB	VNB	VNB	VNB
Study number	9	10	11	12
Author	Hillner & Smith[63]	Smith *et al.*[62]	Evans & Le Chevalier[64]	Vergnenegre *et al.*[66]
Publication year	1996	1995	1996	1996
Base year prices	?	1994/95 ?	1993	1993
Intervention	VNB: 30 mg/m^2/week (iv) VNB+CDDP: VNB 30 mg/m^2/week (iv) and CDDP 120 mg/m^2 on days 1 and 29, then once every 6 weeks VDS+CDDP: VDS 3 mg/m^2/week (iv) for 7 weeks, then once every 2 weeks, and CDDP 120 mg/m^2 on days 1 and 29, then once every 6 weeks	VNB: 30 mg/m^2/week (iv) VNB+CDDP: VNB 30 mg/m^2/week (iv) and CDDP 120 mg/m^2 on days 1 and 29, then once every 6 weeks VDS+CDDP: VDS 3 mg/m^2/week (iv) for 7 weeks, then once every 2 weeks, and CDDP 120 mg/m^2 on days 1 and 29, then once every 6 weeks	VNB: 30 mg/m^2 VNB+CDDP: VNB 30 mg/m^2 and CDDP 120 mg/m^2 VDS+CDDP: VDS 3 mg/m^2 and CDDP 120 mg/m^2 VP-16+CDDP: VP-16 100 mg/m^2 and CDDP 25 mg/m^2 VBL+CDDP: VBL 5 mg/m^2 and CDDP 100 mg/m^2 BSC	MITO+VNB+CDDP: MITO 8 mg/m^2 on days 1, 29 and 71; VNB 25 mg/m^2/week for 16 weeks; and CDDP 120 mg/m^2 on days 1, 29 and 71 MITO+VDS+CDDP: MITO 8 mg/m^2 on days 1, 29 and 71; VDS 3 mg/m^2 on days 1, 8, 15, 22 and 29 and then every fortnight; and CDDP 120 mg/m^2 on days 1, 29 and 71
Study type	CEA model	CEA model	CEA – population health model	CEA model (decision analysis)
NSCLC stage	III and IV	III and IV	IV	III and IV
Perspective	Societal	Societal	Government as payer	Direct hospital costs
Industry role	Collaborator	Part funding through unrestricted grant	Indirect funder, but no restraints placed	None ?
Country of origin	USA	USA	Canada	France

continued

	VNB	PAX	PAX	GEM, VNB and PAX
Study number	13	14	15	16
Author	Evans[72]	Annemans *et al.*[68]	Earle & Evans[89]	Berthelot *et al.*[54]
Publication year	1998	1999	1997	2000
Base year prices	1993	?	1993, 1995	1995
Intervention	VNB: 30 mg/m^2 (iv) weekly VNB+CDDP: VNB 30 mg/m^2 (iv) weekly and CDDP 120 mg/m^2 on days 1 and 29, then every 6 weeks VDS+CDDP: VDS 3 mg/m^2 weekly for 7 weeks, then every 2 weeks, and CDDP 120 mg/m^2 on days 1 and 29, then every 6 weeks VP-16+CDDP: VP-16 100 mg/m^2 (iv) on days 1–3 and CDDP 25 mg/m^2 on days 1–3 every 3 weeks VBL+CDDP: VBL 5 mg/m^2 (iv) on days 1 and 8, and CDDP 100 mg/m^2 (iv) on day 1 every 4 weeks BSC	PAX+CDDP: PAX 175 mg/m^2 on day 1 and CDDP 80 mg/m^2 on day 1 VM-26+CDDP: VM-26 100 mg/m^{2}on days 1, 3 and 5, and CDDP 80 mg/m^2 on day 1	PAX: 214 mg/m^2 (iv) every 3 weeks BSC	VDS+CDDP: VDS 3 mg/m^2 weekly for 7 weeks, then every 2 weeks, and CDDP 120 mg/m^2 on days 1 and 29 every 6 weeks VP-16+CDDP: VP-16 100 mg/m^2 (iv) on days 1–3 every 3 weeks and CDDP 25 mg/m^2 (iv) on days 1–3 every 3 weeks VBL+CDDP: VBL 5 mg/m^2 (iv) on days 1 and 8 every 4 weeks, and CDDP 100 mg/m^2 (iv) on day 1 every 4 weeks GEM: 1000 mg/m^2 weekly x 3 every 4 weeks VNB: 30 mg/m^2 (iv) weekly VNB+CDDP: VNB 30 mg/m^2 (iv) weekly and CDDP 120 mg/m^2 on days 1 and 29 every 6 weeks PAX+CDDP: PAX 135 mg/m^2 (iv) every 3 weeks and CDDP 75 mg/m^2 every 3 weeks PAX+CDDP: PAX 200 mg/m^2 every 3 weeks and CDDP 75 mg/m^2 every 3 weeks PAX+CDDP: PAX 250 mg/m^2 every 3 weeks and CDDP 75 mg/m^2 every 3 weeks
Study type	CEA model	CEA model	Cost minimisation/CEA	CEA/cost–utility analysis
NSCLC stage	IV	IV ?	IV	IV ?
Perspective	Government as payer of universal healthcare system	Healthcare payer	Government as payer	Government as payer
Industry role	Funder – acknowledgement	Funder	Funder	None stated
Country of origin	Canada	Belgium	Canada	Canada

?, unclear information reported; CEA, cost-effectiveness analysis

Appendix 11

Summary of cost-effectiveness results

Results	Interpretation – study's conclusions
***Study 1: Copley-Merriman* et al.*, 1996*[56]** **Base case** Designed as a cost-minimisation study. Both arms assumed to be equally effective. Excluding drug costs, GEM was less expensive by a median of US$2154 (range, US$1504–7425) per cycle compared with VP-16+CDDP **Sensitivity analysis** • Carried out to assess most sensitive variables • Most sensitive variables were treatment of febrile neutropenia and number of days of hospitalisation for chemotherapy administration. Severe febrile neutropenia in VP-16+CDDP arm increased cost savings by 171%	**Conclusion** Single-agent GEM is cost-saving compared with VP-16+CDDP **Caveats** • Assumed equal efficacy • Hospital payments/employer insurance claims databases • Mean cycles: 3.5 for GEM, 2.8 for VP-16+CDDP • Drug costs excluded • Modelling using decision tree analysis • Most cost savings due to outpatient instead of inpatient drug administration • US community care setting • US costing but European, Canadian and US efficacy data • No QoL data • Patient/carer/family costs/benefits not shown
***Study 2: Copley-Merriman* et al.*, 1996*[57]** **Base case** Cost savings associated with GEM arms in all three centres (including Studies 1 and 6 here) varied between US$892 and US$1879 per chemotherapy cycle, compared with VP-16+CDDP or IFOS+VP-16. Savings attributed to reduced hospitalisation for administration and treatment for toxicity **Sensitivity analysis** • Most sensitive variables were number of days of hospitalisation and treatment of febrile neutropenia. Severe febrile neutropenia in VP-16+CDDP arm increased cost savings by 171% • Most significant sources of savings were location of drug administration, treatment of febrile neutropenia and antiemetics use	**Conclusion** Excluding drug costs and assuming equal efficacy, single-agent GEM appears to offer cost savings over the alternative chemotherapy regimens **Caveats** • Assumed equal efficacy • Drug regimen costs excluded (assumed to be equal) • No QoL data
***Study 3: Evans, 1997*[55]** **Base case** Small incremental cost of GEM over BSC regimen. When BSC is used as the base case, assuming survival improvement of 0.4 years with GEM (from BR5 study), GEM cost per LYG (assuming efficacy from EO-18 RCT), at various doses per treatment: 1000 mg/m^2, Can$632 1250 mg/m^2, Can$2796 1500 mg/m^2, Can$4958 1800 mg/m^2, Can$7555 2000 mg/m^2, Can$9285 **Sensitivity analysis** • Survival of GEM patients was reduced by 25% and 50% for each dose per treatment level. ICER ranged from Can$919 to Can$11,782 for 25% reduction in survival and between Can$1578 and Can$17,390 for 50% reduction in survival • Hospitalisation days of GEM patients (on 1000 mg/m^2 dose) for terminal care also varied between 14.1 and 23 days, giving ICER of between –Can$4257 and Can$10,249, cost per LYG	**Conclusion** At low doses, GEM appears to offer a reasonable cost per LYG over the Canadian BSC practice. Cost-effectiveness was stated to be competitive with many other commonly accepted healthcare practices **Caveats** • Canadian population model – microsimulation • Model may underestimate total cost of care, therefore underestimating the cost-effectiveness of GEM • Cost of treatment-related toxicities was not included in model. Authors cited good tolerance profile as justification • Model overestimates impact of GEM on the total healthcare budget in Canada • Calculation assumes mean of 3.3 cycles of GEM • Patient/family/social benefits/costs not included • No QoL data • Cost of diagnosis included • Follow-up cost included • Authors stated terminal-care hospitalisation for GEM was estimated as equated to other regimens – no reliable estimate was available

continued

Results	Interpretation – study's conclusions
Study 4: Evans, 1996[61] **Base case** Cost of BSC was estimated at Can$20,914, GEM at Can$22,172 (based on 3.3 cycles at Can$1000 per cycle). Therefore, incremental cost is Can$1258 With BSC used as the base case, GEM cost per LYG (at various drug costs per treatment cycle): Can$1609 (Can$800 per cycle) Can$3193 (Can$1000 per cycle) Can$4777 (Can$1200 per cycle) Can$9529 (Can$1800 per cycle) **Sensitivity analysis** • Survival of GEM patients was reduced by 25% and 50% for each cost per treatment cycle. ICER ranged from Can$2054 to Can$11,956 for 25% reduction in survival and between Can$2839 and Can$16,230 for 50% reduction in survival • Hospitalisation days of GEM patients (based on Can$1000 cost per treatment cycle) for terminal care were also varied, giving ICER of Can$1002, cost per LYG, for each additional day of hospitalisation	**Conclusion** In terms of cost-effectiveness, it appears that GEM chemotherapy for stage IV NSCLC is competitive with many other commonly accepted healthcare practices. In terms of a cost of Can$1000 per treatment cycle, the cost per LYG is Can$3193 and is very cost-effective **Caveats** • Dual publication? Very similar to Study 3 (Evans, 1997[55]) • Population model – microsimulation • Model may underestimate total cost of care, therefore underestimating the cost-effectiveness of GEM • Cost of treatment-related toxicities was not included in model • Model overestimates impact of GEM on the total healthcare budget in Canada, because not all stage IV NSCLC patients will be treated with GEM • Patient/family/social benefits/costs not included • QoL not included • Canadian model • Cost of diagnosis included • Follow-up cost included
Study 5: Palmer & Brandt, 1996[59] **Base case** Average cost-effectiveness analysis showed no significant difference between treatments. Marginal CEA showed the use of MITO+IFOS+CDDP, VP-16+CDDP or VNB+CDDP instead of GEM+CDDP would result in additional costs of 7.7, 55.2 and 46.2 million lira, respectively, for every patient with a tumour response. (Response rates used were 54% for GEM+CDDP, 40% for MITO+IFOS+CDDP, 26% for VP-16+CDDP and 35% for VNB+CDDP) **Sensitivity analysis** • Effectiveness varied within bounds of their 95% CI • One-way sensitivity analysis was performed to find greatest influential factors • Tumour response was most influential factor, followed by neutropenia and anaemia • Average cost-effectiveness ranges (million lira): GEM+CDDP, 48.5–147.1 MITO+IFOS+CDDP, 59.2–143.4 VP-16+CDDP, 65.6–253.9 VNB+CDDP, 81.4–225.4	**Conclusion** GEM+CDDP is more effective, to varying degrees of significance, than the other three regimens. Marginal CEA demonstrated an advantage for GEM+CDDP over the other regimens **Caveats** • Effectiveness data were gathered via meta-analysis of other studies that follow Italian perspective • Effectiveness was expressed in terms of cost per response, which may not reflect survival or QoL • Italian DRG costs • Model of costing • QoL data not shown • A lot of cost savings were attributed to shifting from inpatient to outpatient chemotherapy administration
Study 6: Koch* et al., *1995[88] **Base case** Compared with IFOS+VP-16, single-agent GEM was associated with cost savings of DM3026 (DM11,151 for IFOS+VP-16 vs DM5798 for GEM, both calculated based on 2 cycles). This did not include chemotherapy costs, but included 40% savings from hospitalisation and 54% savings in managing toxic effects **Sensitivity analysis** • Reducing cost of outpatient visit to 50% of inpatient resulted in further savings for GEM • Noted that varying other key parameters did not have a significant impact on findings	**Conclusion** Hospital costs (excluding drug costs) would appear to favour GEM over IFOS+VP-16. Most of these savings are due to a shift away from the inpatient setting **Caveats** • Unclear how widely used IFOS+VP-16 combination is – said to be representative of combination chemotherapy in Germany • Costs of chemotherapy may offset other savings • Efficacy assumed equal (based on response rates and median survival) • Cost of follow-up visits not included • No QoL estimates • Calculations assumed IFOS+VP-16 was given on inpatient basis (over 5 days), but stated that it could be administered over less time on an outpatient basis. No costs of this alternative administration approach were provided

continued

Results	Interpretation – study's conclusions
***Study 7: Sacristan* et al., *2000*[60]** **Base case** Assuming equal survival, there was no statistical difference in direct costs of the GEM+CDDP or VP-16+CDDP regimens (only the chemotherapy cost itself was significant). The total direct costs of GEM+CDDP were 584,523 ptas (SD, 281,201 ptas) and of VP-16+CDDP 589,630 ptas (SD, 601,102 ptas). Higher drug costs of GEM+CDDP were offset by reduced hospital stay. Hospitalisation was 4.8 days (SD, 6.2 days) for the GEM+CDDP arm and 9.1 days (SD, 14.2 days) for the VP-16+CDDP arm **Sensitivity analysis** • When the cost of hospitalisation was varied by –25% to +50%, this impacted on direct costs in favour of VP-16+CDDP and GEM+CDDP, respectively • Differences in efficacy were not examined in the sensitivity analysis. Cost-effectiveness and ICER were used when outcomes were measured by response and time to disease progression. Cost per response favoured GEM+CDDP at 1,439,712 ptas vs 2,692,374 ptas with VP-16+CDDP. Cost per progression-free month also favoured GEM+CDDP at 84,713 ptas vs 137,123 ptas with VP-16+CDDP. ICER favoured GEM+CDDP: incremental cost-effectiveness savings per additional response were 27,310 ptas and incremental cost-effectiveness savings per progression-free month were 3405 ptas • CIs for ICER were examined using bootstrapping (25,000 iterations) in which most points were clustered in an ellipse around the positive x-axis, indicating that efficacy gains with GEM will be accompanied by small differences in cost	**Conclusion** The two regimens are of roughly equivalent average cost because the increased drug cost in the GEM arm was offset by reduced hospitalisation **Caveats** • Differences in hospitalisation, although larger (which offset the higher GEM drug cost), were not statistically significant; therefore, results should be viewed with caution • Efficacy taken from Phase III clinical trial (Cardenal *et al.*, 1999[26]) – no difference in survival between regimens • Superiority of GEM+CDDP in terms of response and progression of disease may impact on QoL, even though efficacy is unaffected • No QoL estimates
***Study 8: Tennvall & Fernberg, 1998*[58]** **Base case** Designed as a cost-minimisation study. The total costs of GEM alternative were approximately Skr 1800 and Skr 12,000 lower than those of VP-16+CDDP and IFOS+VP-16, respectively. Results assume GEM was used in outpatients, while other treatments were used in inpatients. If all patients were treated with 3 cycles of chemotherapy, the average cost of a complete treatment with GEM would be Skr 3203 and Skr 30,990 lower than VP-16+CDDP and IFOS+VP-16, respectively **Sensitivity analysis** • If all GEM patients were treated as inpatients, then GEM would still be less costly than IFOS+VP-16 • Setting of treatment was changed to see how sensitive the results are • GEM is always less costly, unless both GEM and VP-16+CDDP are administered 100% in inpatient setting, or both GEM and VP-16+CDDP are administered 100% in outpatient setting	**Conclusion** GEM in outpatient treatment is the least costly **Caveats** • Efficacy assumed to be equivalent for treatment options • No QoL assessments • Patient/family/social benefits/costs not included, except for travel costs incurred by patients • Evaluation has been based on data from different sources, and assumptions have been made • Difficult to compare efficacy between RCTs • Costing of healthcare resources based on Swedish treatment practice

continued

Results	Interpretation – study's conclusions
Study 9: Hillner & Smith, 1996[63] **Base case** Using VNB alone as base case: • VNB+CDDP added 56 days of life at an additional cost of US$2700, giving an ICER of US$17,700 per LYG • VDS+CDDP added 19 days of life at an additional cost of US$1150, giving an ICER of US$22,100 per LYG Using VDS+CDDP as base case: • VNB+CDDP added 37 days of life at an additional cost of US$1570, giving an ICER of US$15,500 per LYG **Sensitivity analysis** • None carried out	**Conclusion** Compared with other medical intervention, it is the authors' opinion that the ICER of VNB+CDDP is most cost-effective compared with both VNB alone and VDS+CDDP **Caveats** • Cost data were collected separately from the trial, using a US medical college. The RCT was performed in Europe • QoL data not shown • Patient/carer/family benefits/costs not shown • Societal costs excluded
Study 10: Smith* et al., *1995[62] **Base case** Using VNB alone as base case: • VNB+CDDP added 56 days of life at an additional cost of US$2700, giving an ICER of US$17,700 per LYG • VDS+CDDP added 19 days of life at an additional cost of US$1150, giving an ICER of US$22,100 per LYG Using VDS+CDDP as base case: • VNB+CDDP added 37 days of life at an additional cost of US$1570, giving an ICER of US$15,500 per LYG **Sensitivity analysis** • If survival benefit is increased, then ICER increases also (survival benefit between baseline and 200%), meaning an ICER between US$8800 and US$34,200 (for VNB +CDDP vs VNB alone), an ICER between US$11,000 and US$59,000 (for VDS+CDDP vs VNB alone), an ICER between US$7700 and US$29,300 (for VNB+CDDP vs VDS+CDDP) • Cost–utility analysis was carried out to adjust for QoL. Using VNB alone as base case, VNB+CDDP added 4.5 days at a cost per LYG of US$241,000. VDS+CDDP gave 18 additional days with cost-savings of US$29,000 per LYG. Using VDS+CDDP as base case, VNB+CDDP added 22 days of life at a cost per LYG of US$25,800 • The cost of CDDP, cost of CDDP administration, length of inpatient stay and cost of toxicity were varied	**Conclusion** Compared with other medical intervention, it is the authors' opinion that the ICER of VNB+CDDP is most cost-effective compared with both VNB alone and VDS+CDDP **Caveats** • Cost data were collected separately from the trial, using a US medical college. The RCT was performed in Europe • No patient perspective on QoL. Utility estimates (0.7 for VNB, 0.6 for CDDP-containing regimens) were based on the experience of 14 oncology physicians and nurses • No patient/carer/family benefits/costs • Societal costs excluded • Costs of treatment vary in different countries

continued

Results	Interpretation – study's conclusions
Study 11: Evans & Le Chevalier, 1996[64] **Base case** VBL+CDDP was the most cost-effective regimen relative to BSC, with an increase in survival of 0.27 years/patient and a reduction in cost of Can$3265 per person. VNB+CDDP achieved a greater increase in survival of 0.44 years, but costs increased by Can$2451 per person. Cost-effectiveness ratio was Can$5551 per LYG. Cost savings were also obtained for single-agent VNB and VNB+CDDP when given on an outpatient basis **Sensitivity analysis** • Survival: A reduction in survival gain by 50% maintained cost savings of VNB and outpatient VNB over BSC • Hospitalisation for terminal care: Assuming equal days for terminal care as with BSC, costs per LYG were Can$9779 for VNB and Can$8420 for outpatient VNB+CDDP, respectively. Inpatient VNB+CDDP with terminal stay of 23.6 days resulted in cost per LYG of Can$15,042	**Conclusion** VBL+CDDP is cost-saving compared with BSC but provides a smaller survival gain than VNB+CDDP. VNB+CDDP achieves greatest survival gains but at an increased cost over BSC **Caveats** • Data were mixed from European and Canadian trials: some chemotherapy regimens were taken from European trial, but administration costs were Canadian; survival taken from European trial (Le Chevalier *et al.*, 1994[74]) for some regimens, Canadian BR5 trial (NCIC) for others; survival with VNB+CDDP (and VP-16) was assumed equal to that with VDS+CDDP, based on previous trials • Hospitalisation cost was taken from BR5 trial update from 1984 prices to 1993; VNB+CDDP administration cost was estimated from Canadian practice • Used Canadian POHEM model • Cost of BSC taken from Canadian BR5 trial and adjusted to 1993 prices • Duration of hospitalisation for terminal care for chemotherapy regimens was based on 17.1 days vs 23.6 with BSC, based on 1984 BR5 trial • No account was taken of the cost of complications or toxicity with chemotherapy regimens
Study 12: Vergnenegre* et al., *1996[66] **Base case** MITO+VNB+CDDP produced a benefit of 12339.40 FF per response over MITO+VDS+CDDP, based on response rates of 25% and 17%, respectively **Sensitivity analysis** • Cost and response rates were varied to see the effect on MITO+VNB+CDDP • By varying the cost by 20% around the observed value, most points (average cost-effectiveness ratio of MITO+VNB+CDDP with respect to MITO+VDS+CDDP) were found to lie in the quadrant of medical benefit and decrease in cost • A benefit for MITO+VDS+CDDP, using the values observed, occurred only by a considerable increase in MITO+VNB+CDDP costs with an effectiveness below 0.20 • An analysis of the sensitivity on response rates of the two strategies assumed constant costs • The excess costs of MITO+VNB+CDDP and the benefits of MITO+VNB+CDDP are shown as a function of the response rate in a figure in the paper. Apart from a response rate of 14%, varying degrees of benefit emerge consistently for MITO+VNB+CDDP	**Conclusion** Cost analysis of hospital administration of the two chemotherapeutic regimens showed that MITO+VNB+CDDP has a low increase of average direct costs per patient. MITO+VNB+CDDP has a more favourable cost-effectiveness ratio **Caveats** • Effectiveness was measured in terms of response • Results were obtained from an RCT and only valid for this group of patients • Direct costs were collected from hospital perspective • No account of indirect costs • French accounting and care methods were used • Incremental cost-effectiveness not calculated • Societal costs excluded • Patient/carer/family cost/benefits not shown • QoL not shown

continued

Results	Interpretation – study's conclusions
Study 13: Evans, 1998[72] **Base case** Using BSC as base case: • VNB alone added 0.29 average years of life at a cost saving of Can$1014, thus strategy is cost-dominant • VNB+CDDP (outpatient) added 0.45 average years of life at a cost saving of Can$22, thus strategy is cost-dominant • VNB+CDDP (inpatient) added 0.45 average years of life at an ICER of Can$6386 per LYG • VDS+CDDP added 0.27 average years of life at an ICER of Can$13,568 per LYG • VP-16+CDDP added 0.27 average years of life at a cost saving of Can$1418, thus strategy is cost-dominant • VBL+CDDP added 0.27 average years of life at a cost saving of Can$2865, thus strategy is cost-dominant Using VP-16+CDDP as base case: • VNB alone gave a non-significant ICER per LYG • VNB+CDDP (outpatient) gave an ICER of Can$7450 per LYG • VNB+CDDP (inpatient) gave an ICER of Can$23,053 per LYG • VDS+CDDP gave a non-significant ICER per LYG Using VBL+CDDP as base case: • VNB alone gave a non-significant ICER per LYG • VNB+CDDP (outpatient) gave an ICER of Can$15,171 per LYG • VNB+CDDP (inpatient) gave an ICER of Can$30,770 per LYG • VDS+CDDP gave a non-significant ICER per LYG **Sensitivity analysis** • Sensitivity analysis reduced survival gain of VNB+CDDP patients by 25% and 50%. A 50% reduction in survival gain led to a cost per LYG for outpatient VNB+CDDP of Can$443 compared with BSC • Sensitivity analysis also increased the number of hospital bed-days for terminal care from 17.1 to 23.6 in the chemotherapy-treated patients. Cost per LYG for the outpatient VNB+CDDP arm was Can$9195 when compared with BSC for 23.6 days of terminal care	**Conclusion** When BSC is used as the base case, then VBL+CDDP and VP-16+CDDP are the most cost-dominant, whereas VNB+CDDP (both inpatient and outpatient) has the greater average LYG. Outpatient VNB+CDDP is cost-dominant, while inpatient VNB+CDDP has an ICER of Can$6386. When VP-16+CDDP is used as the base case, outpatient VNB+CDDP is most cost-effective at Can$7450. When VBL+CDDP is used as the base case, outpatient VNB+CDDP again is most cost-effective at Can$15,171 **Caveats** • Not all complications and treatments were accounted for • Outside costs were not accounted for – home care visits, family physician involvement in palliative care, terminal care, indirect costs incurred by patients in terms of travelling • No QoL estimates • Efficacy data were collected from other RCT conducted in Europe, Canada and other countries
Study 14: Annemans* et al., *1999[68] **Base case** Using VM-26+CDDP as base case: • ICER per responder for PAX+CDDP for different countries: The Netherlands, US$33,550; Belgium, US$17,775; Spain, US$20,394; France, US$12,323 • Cost per responder was US$21,011 for PAX+CDDP compared with US$27,266 for the VM-26+CDDP regimen (based on response rates of 37% and 26%, respectively) **Sensitivity analysis** • Probability of hospitalisation for haematological side-effects varied, and G-CSF use varied (results shown in graph) • Cost of acquiring PAX varied by 30% each way. ICER ranged from –US$4992 to US$56,931 • Hospital costs varied by 30% each way. ICER was between US$1753 and US$40,023 • Use of different databases was compared (EORTC or Bristol-Myers Squibb). ICER was between US$12,323 and US$33,550 • Cost per physical functioning score point was generated from EORTC QLQ-C30 questionnaire (average score was 47/100 for VM-26+CDDP group vs 75/100 for PAX+CDDP group). Cost-effectiveness scores and ICERs favoured PAX+CDDP	**Conclusion** PAX+CDDP does not offer additional survival in NSCLC patients but improves significantly patient response and some dimensions of QoL. Expressed as cost per responder, PAX+CDDP is equally cost-effective as VM-26+CDDP. Analysis suggests that PAX+CDDP can be considered as a cost-effective intervention **Caveats** • Costs used in the study were charges not 'real' costs • Labour time in management of patients was not calculated • Delphi approach was used to collect data regarding medical practice • Cost-effectiveness was restricted to responders to treatment • Limited QoL estimates were based on physical functioning scale points • Costs altered to US$ • Costing and treatment set-up in European countries

continued

Results	Interpretation – study's conclusions
Study 15: Earle & Evans, 1997[89] **Base case** Using BSC as the base case, the incremental cost of single-agent PAX was Can$3375, or a total cost of Can$8143 per patient based on 3 cycles of chemotherapy. Cost per LYS (based on survival from a Phase II trial at 7.9 months more than BSC) was Can$4778 **Sensitivity analysis** • Days of hospitalisation for terminal care (increased from 17 to 23 days) led to a cost per LYS of Can$10,519 • Number of chemotherapy cycles increased to 4 or 5; 5 cycles of chemotherapy led to a cost per LYS of Can$10,788 • Survival reduced by 25–50%; a 50% reduction in survival led to Can$9757 • Worst reported case was cost per LYS of Can$21,377, with both a 50% reduction in increased survival and 5 cycles of chemotherapy	**Conclusion** PAX appears to offer relatively favourable cost per LYS and incremental cost compared with BSC. However, the calculations do not take into account any QoL effects. Length of terminal stay also needs to be affirmed for PAX **Caveats** • Mean PAX dose was 214 mg/m^2 • Worst reported case in sensitivity analysis was only multiway sensitivity scenario. Could also have added increased hospitalisation for terminal care, as was considered in one-way analysis. • Cost of BSC in these studies was calculated from an RCT • Hospitalisation for terminal care was based on BR5 regimens; PAX just assumed to be equivalent due to similar response rates • No QoL estimates • Average survival duration was based on patients from two separate studies and only 49 patients • Confusion over base year for costs (1993 or 1995?) • Cost of hospitalisation was taken from BR5 trial and staff costs from Ottawa Cancer Centre • Study was funded by grant from drug company
Study 16: Berthelot* et al., *2000[54] **Base case** Regimens were compared with BSC cost (Can$25,904). One regimen, VBL+CDDP, yielded longer survival at a lower cost than BSC estimates (average cost, Can$24,828) Incremental costs per LYS (compared with BSC) were Can$1900 for VNB, Can$4100 for VNB+CDDP, Can$6800 for GEM, Can$15,400 for PAX (135 mg), Can$21,500 for PAX (200 mg) and Can$27,000 for PAX (250 mg) VBL+CDDP was the only dominant strategy compared with BSC when cost per QALY gained was considered. Other cost per QALYs (vs BSC) were Can$2700 for VNB, Can$6000 for VNB+CDDP, Can$8600 for GEM, Can$21,500 for PAX (135 mg), Can$30,100 for PAX (200 mg) and Can$37,800 for PAX (250 mg) **Sensitivity analysis** • Survival and number of days of hospitalisation for terminal care were varied in one-way analysis • VBL+CDDP remained dominant to BSC, with 25% and 50% reductions in survival gain. With an increase in hospitalisation for terminal care from 17 to 23 days, cost per LYG was Can$12,200 compared with BSC. This illustrates the sensitivity to hospitalisation	**Conclusion** The authors reported that strategy of choice depends on an acceptable cost-effectiveness threshold. VBL+CDDP is the lowest cost option and the regimen of choice up to a threshold valve of up to Can$10,000 cost per LYS or cost per QALY. VNB or VNB+CDDP is their choice when a cost per LYS threshold of Can$25,000–50,000 is acceptable. With a threshold of Can$75,000–100,000, PAX(135)+CDDP would be the preferred regimen. Taking into account their QoL utility estimates, GEM would be preferable using a Can$50,000 threshold. The authors recommend the use of chemotherapeutic regimens and the abandonment of BSC **Caveats** • Utility estimates used in QALY calculations were derived by 24 oncologists • Mean survival data were based on single Phase III studies (Phase II in case of GEM) and recalculated from raw data • Average BSA used was 1.73 m^2 • Produced tables and graphs of cost-effectiveness ranking depending on acceptable threshold of cost per LYS and cost per QALY • Cost and efficacy data were mixed from European and North American trials and Canadian practice • Transparency uncertain because used Canadian POHEM model • Duration of hospitalisation for terminal care for chemotherapy regimens was based on 17 days vs 23 days with BSC, based on 1984 BR5 trial • No account was taken of cost of complications or toxicity with chemotherapy regimens

ICER, incremental cost-effectiveness ratio; DRG, diagnosis-related group; DM, Deutschmark; ptas, pesetas; Skr, Swedish krona; FF, French francs

Appendix 12

Internal validity of economic evaluations

Item	GEM	GEM	GEM	GEM
	Study 1: Copley-Merriman *et al.*, 1996[56]	**Study 2: Copley-Merriman *et al.*, 1996**[57]	**Study 3: Evans, 1997**[55]	**Study 4: Evans, 1996**[61]
1. Well-defined question	✔	✔	✔	✔
2. Clear description of alternatives	✔ GEM, VP-16+CDDP	✔ GEM, VP-16+CDDP, IFOS+VP-16	✔ GEM, BSC	✔ GEM, BSC
3. Reasonable study type	✔	✔ Three multinational healthcare settings	✔	✔
4. Effectiveness established	Efficacy assumed to be equal	Efficacy assumed to be equal	✔ Effectiveness (survival data) established from international RCT EO-18 (GEM) and Canadian RCT BR5 (BSC)	✔ Effectiveness (survival data) established from international RCT EO-18 (GEM) and Canadian RCT BR5 (BSC)
5. Estimates related to population risks	?	?	?	?
6. Relevant costs and consequences identified	✔ Healthcare resources ✔ ✘ Adverse events (only high cost included) ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✔ ✘ Adverse events (only high cost included) ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ Adverse effects ✔ Drug costs ✔ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ Adverse effects ✔ Drug costs ✔ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits
7. Costs and consequences measured accurately	✔ Measured in appropriate physical units	✔ Costs measured differently in the three models, depending on the amount of data available for each country. Only direct costs calculated. No account of patient QoL associated with toxicity	✔ Resource use stated, costs measured in appropriate physical units	✔ Resource use stated, costs measured in appropriate physical units
8. Costs and consequences valued credibly	✔ ? Included all direct treatment costs; included only high-cost toxic events	✔ ✘ Costs valued credibly, consequences not	✔ Direct costs included, no indirect costs	✔ Direct cost included, no indirect costs
9. Differential timing considered	✘	✘	✘	✘
10. Incremental analysis performed	✔	✔	✔	✔
11. Sensitivity analysis performed	✔	✔ Performed in all three models	✔	✔
12. Modelling conducted reasonably	?	?	?	?

continued

Item	GEM and VNB	GEM	GEM	GEM
	Study 5: Palmer & Brandt, 1996[59]	**Study 6: Koch *et al.*, 1995[88]**	**Study 7: Sacristan *et al.*, 2000[60]**	**Study 8: Tennvall & Fernberg, 1998[58]**
1. Well-defined question	✔	✔	✔	✔
2. Clear description of alternatives	✔ GEM+CDDP, MITO+IFOS+CDDP, VP-16+CDDP, VNB+CDDP	✔ GEM, VP-16+IFOS	✔ GEM+CDDP, VP-16+CDDP	✔ GEM, VP-16+CDDP, IFOS+VP-16
3. Reasonable study type	✔	✔ VP-16+IFOS representative of combination chemo-therapy in Germany	✔	✔ Given that efficacy is assumed to be equivalent for all interventions
4. Effectiveness established	✔ Efficacy and safety data were calculated from studies selected from international literature, using formal inclusion and exclusion criteria	✔ Separately published multicentre, Phase II RCT for GEM ✘ No RCTs cited for VP-16+IFOS efficacy	✔ Separately published multicentre, Phase II RCT	Efficacy (survival and tumour response) assumed to be equivalent for all treatment arms
5. Estimates related to population risks	?	?	?	?
6. Relevant costs and consequences identified	✔ Healthcare resources ✔ Drug costs ✔ ✘ Adverse events (only WHO grade 3 or 4) ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ GEM drug cost ✘ Adverse events ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✔ Drug costs ✔ Adverse events ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✔ Adverse events ✔ Drug costs ✔ Follow-up visits ✔ ? Patient/family resources (patient's travelling costs included) ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits
7. Costs and consequences measured accurately	✔ Measured from government reimburse-ment perspective based on DRGs	✔ Measured in appropriate physical units	✔ Focus on high-cost resources – patient-collected data from trial, unitary costs from govern-ment and hospital data	✔ Measured in appropriate physical units
8. Costs and consequences valued credibly	✔ Included all direct medical costs and management of WHO grade 3 and 4 toxicity from government re-imbursement perspective	✔ Except outpatient cost given as equal to inpatient cost	✔ Direct costs only. Pesetas converted to 1997 US$	✔ All direct costs as well as side-effects, but no indirect costs, except for travelling costs of patients
9. Differential timing considered	✘	✘	✘	✘
10. Incremental analysis performed	✔ Average and marginal CEA performed	✔ But did not include GEM drug cost	✔ Response and progression parameters	✘ Was not CEA. Total direct costs per patient established
11. Sensitivity analysis performed	✔ One-way sensitivity analysis of average cost-effectiveness; each parameter varied at one time. Cost varied by 10%, efficacy and toxicity varied by 95% CI. "Analysis of extremes" carried out	✔ One-way sensitivity analysis; unclear how many variables ✘ IFOS+VP-16 administered on inpatient basis but stated can also be given on outpatient basis. However, outpatient cost not considered	✔ One-way sensitivity analysis for hospitalisation only; bootstrapping conducted around CIs for cost and efficacy (response and progression delay)	✔ Changing proportion of inpatient and outpatients. Patient travel distance was varied too
12. Modelling conducted reasonably	?			?

continued

Item	VNB	VNB	VNB	VNB
	Study 9: Hillner & Smith, 1996[63]	**Study 10: Smith *et al.*, 1995[62]**	**Study 11: Evans & Le Chevalier, 1996[64]**	**Study 12: Vergnenegre *et al.*, 1996[66]**
1. Well-defined question	✔	✔	✔	✔
2. Clear description of alternatives	✔ VNB, VNB+CDDP, VDS+CDDP	✔ VNB, VNB+CDDP, VDS+CDDP	✔ VNB, VNB+CDDP, VDS+CDDP, VP-16+CDDP, VBL+CDDP, BSC	✔ VNB+MITO+CDDP, VDS+MITO+CDDP
3. Reasonable study type	✔	✔	✔	✔
4. Effectiveness established	✔ Established in an RCT published separately	✔ Established in an RCT published separately	✔ Established in previous RCTs	✔ Effectiveness established in an RCT, described in the paper
5. Estimates related to population risks	?	?	?	?
6. Relevant costs and consequences identified	✔ Healthcare resources but ? adverse events ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources but ? adverse events ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ Adverse events ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✔ Adverse effects ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits
7. Costs and consequences measured accurately	✔ Measured in appropriate physical units	✔ Measured in appropriate physical units	✔	✔ Measured in appropriate physical units
8. Costs and consequences valued credibly	✔ Included all direct treatment costs; included toxic effects and monitoring performed as though all patients had same supportive care	✔ Included all direct treatment costs; included toxic effects and monitoring performed as though all patients had same supportive care	✔ Direct costs only	✔ Included all direct costs as well as side-effects
9. Differential timing considered	✘	✘	✘	✘
10. Incremental analysis performed	✔	✔	✔	✘ Average cost-effectiveness established
11. Sensitivity analysis performed	✘	✔ Incremental survival benefit and costs compared with another common US treatment regimen	✔ One-way	✔ Two types of sensitivity analysis performed: (1) sensitivity of one strategy on cost and response rates (2) sensitivity on response rates of the two strategies, assuming constant costs
12. Modelling conducted reasonably	?	?	?	?

continued

Item	VNB	PAX	PAX	GEM, VNB and PAX
	Study 13: Evans, 1998[72]	**Study 14: Annemans *et al.*, 1999[68]**	**Study 15: Earle & Evans, 1997[89]**	**Study 16: Berthelot *et al.*, 2000[54]**
1. Well-defined question	✔	✔	✔	✔
2. Clear description of alternatives	✔ VNB, VNB+CDDP, VP-16+CDDP, VDS+CDDP, VBL+CDDP, BSC	✔ VNB, VNB+CDDP, PAX+CDDP, VM-26+CDDP	✔ PAX, BSC	✔ VDS+CDDP, VP-16+CDDP, VBL+CDDP, GEM, VNB, VNB+CDDP, PAX+CDDP, BSC
3. Reasonable study type	✔	✔	✔	✔
4. Effectiveness established	✔ Effectiveness based on other RCT	✔ Established in a separate multicentre European RCT	✔ Pooled two published Phase II trials	✔ Data from Phase III and Phase II trials
5. Estimates related to population risks	?	?	?	?
6. Relevant costs and consequences identified	✔ Healthcare resources ✘ Adverse effects ✔ Drug costs ✔ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✔ Adverse effects ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ Adverse effects ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits	✔ Healthcare resources ✘ Adverse effects ✔ Drug costs ✘ Follow-up visits ✘ Patient/family resources ✘ Social care sector resources ✘ Patient benefits ✘ Carer benefits
7. Costs and consequences measured accurately	✔ Measured in appropriate physical units	? Drug use stated; however, other resource use available on request, otherwise costs measured in appropriate physical units; resource use collected through a Delphi technique from nine randomly selected clinicians	✔ Measured in appropriate physical units	✔ Measured in appropriate physical units
8. Costs and consequences valued credibly	✔ Included all direct costs as well as follow-ups, but not side-effects	✔ ? Direct costs included, but no indirect costs. All costs translated into US$. Charges used not costs – but may be justified because the perspective is a healthcare system that would face charges not costs?	✔ Direct costs only. Used physician fees charged for assessment not costs, also pretreatment fee-based costs. Nursing and pharmacy at cost	✔ Direct costs. Indirect costs not included
9. Differential timing considered	✘	✘	✘	✘
10. Incremental analysis performed	✔	✔	✔	✔
11. Sensitivity analysis performed	✔ Sensitivity analysis for survival performed. Survival gain and terminal care days varied	✔ Medical practice varied. Cost of hospitalisation and acquisition cost of PAX varied by 30% either way. Varied the choice of database between EORTC and Bristol-Myers Squibb	✔	✔ Survival and days hospitalisation for terminal care were varied
12. Modelling conducted reasonably	?	?	?	?

?, unclear or unknown; ✔, item included or judged as acceptable to be internally valid; ✘, factor not included or judged as unacceptable to be internally valid

Appendix 13

External validity of economic evaluations

Item	GEM	GEM	GEM	GEM
	Study 1: Copley-Merriman *et al.*, 1996[56]	**Study 2: Copley-Merriman *et al.*, 1996[57]**	**Study 3: Evans, 1997[55]**	**Study 4: Evans, 1996[61]**
1. Patient group – are the patients in the study similar to those of interest in England and Wales?	? Efficacy obtained from studies carried out in Europe, Canada and USA	? Efficacy obtained from studies carried out in Europe, Canada and USA	? Patient setting was from an international trial and from a Canadian trial	? Patient setting was from an international trial and from a Canadian trial
2. Healthcare system/ setting – comparability of available alternatives? Similar levels of resources? No untoward supply constraints? Institutional arrangements comparable?	✘ US private insurance payer perspective	✘ US private insurance payer, Spanish public healthcare payer, German health insurance perspectives	✘ Canadian healthcare setting, Canadian model	✘ Canadian healthcare setting, Canadian model
3. Treatment – comparability with clinical management?	? Treatment in USA, Canada and European centres	? Treatment in German, Spanish and US centres, including expert opinion	? Treatment was based on international and Canadian trial protocols	? Treatment was based on international and Canadian trial protocols
4. Resource costs – comparability between study and setting/ population of interest?	✘ US cost data	✘ German, Spanish and US cost data	✘ Canadian cost data	✘ Canadian cost data
5. Marginal versus average costs – what difference does this make? Are there real cost savings from averting short periods in long-stay care?	✘	✘	✘	✘

continued

Item	GEM and VNB	GEM	GEM	GEM
	Study 5: Palmer & Brandt, 1996[59]	**Study 6: Koch *et al.*, 1995[88]**	**Study 7: Sacristan *et al.*, 2000[60]**	**Study 8: Tennvall & Fernberg, 1998[58]**
1. Patient group – are the patients in the study similar to those of interest in England and Wales?	? Efficacy data gathered from meta-analysis based on international studies that reflect Italian situation	? Efficacy for GEM from European trial; IFOS+VP-16 based on German and US trials	? Efficacy data from Spanish RCT	? Efficacy data from several different sources, such as clinical trial reports, review articles, local information and trial reports
2. Healthcare system/ setting – comparability of available alternatives? Similar levels of resources? No untoward supply constraints? Institutional arrangements comparable?	✘ Italian third-party (insurance) payer perspective	✘ German healthcare perspective	✘ Spanish healthcare payer	✘ Spanish societal perspective
3. Treatment – comparability with clinical management?	? Reflects current Italian treatment	? German clinical practice	? Treatment from Spanish perspective	? Treatment practice and resource use according to clinical report from Eli Lilly and oncologists; assumptions also made
4. Resource costs – comparability between study and setting/ population of interest?	✘ Italian costs based on Italian DRGs	✘ German cost data	✘ Spanish government and hospital costs	✘ ? Costing based on Swedish costing practice from University Hospital, Lund, Sweden. Drug costs are based on 1996 market prices
5. Marginal versus average costs – what difference does this make? Are there real cost savings from averting short periods in long-stay care?	✘	✘	✘	✘

continued

Item	VNB	VNB	VNB	VNB
	Study 9: Hillner & Smith, 1996[63]	**Study 10: Smith *et al.*, 1995[62]**	**Study 11: Evans & Le Chevalier, 1996[64]**	**Study 12: Vergnenegre *et al.*, 1996[66]**
1. Patient group – are the patients in the study similar to those of interest in England and Wales?	? Efficacy data from an RCT conducted in Europe, but patient characteristics not given	? Efficacy data from an RCT conducted in Europe, but patient characteristics not given	? Efficacy data from RCTs in USA and Canada	? Trial carried out in France
2. Healthcare system/ setting – comparability of available alternatives? Similar levels of resources? No untoward supply constraints? Institutional arrangements comparable?	✘ US healthcare setting	✘ US healthcare setting	✘ Canadian healthcare system	✘ Setting based on French experience
3. Treatment – comparability with clinical management?	? Treatment in European centres	? Treatment in European centres	? Treatment in centres in European trial	? Treatment carried out in various centres in France
4. Resource costs – comparability between study and setting/ population of interest?	✘ US costing data	✘ US costing data	✘ Canadian costs	✘ Costing carried out in France according to the methodology of a Canadian study – based on two hospitals
5. Marginal versus average costs – what difference does this make? Are there real cost savings from averting short periods in long-stay care?	✘	✘	✘	✘

continued

Item	VNB	PAX	PAX	GEM, VNB and PAX
	Study 13: Evans, 1998[72]	**Study 14: Annemans *et al.*, 1999[68]**	**Study 15: Earle & Evans, 1997[89]**	**Study 16: Berthelot *et al.*, 2000[54]**
1. Patient group – are the patients in the study similar to those of interest in England and Wales?	? Efficacy data from RCTs carried out in Europe and Canada	? Patient setting was from European trial	? Efficacy data from two Canadian trials	? Efficacy data from North American and European trials
2. Healthcare system/ setting – comparability of available alternatives? Similar levels of resources? No untoward supply constraints? Institutional arrangements comparable?	? Setting based on Canadian experience (government payer in universal healthcare system)	✘ Insurance system in each of four countries	? Setting based on Canadian experience (government payer in universal healthcare system)	? Setting based on Canadian experience (government payer in universal healthcare system)
3. Treatment – comparability with clinical management?	? Treatment was based on European trial and Canadian experience	? Treatment was based in European centres: The Netherlands, Belgium, Spain and France	? Treatment was average of two Canadian trials	? Treatment was based on North American and European trial experience
4. Resource costs – comparability between study and setting/ population of interest?	✘ Canadian cost data	✘ Costs obtained from each of the four European Union countries	✘ Canadian cost data	✘ Canadian cost data
5. Marginal versus average costs – what difference does this make? Are there real cost savings from averting short periods in long-stay care?	✘	✘	✘	✘

?, unclear or unknown; ✔, judged item suitable to generalise to England and Wales, with or without some readjustment; ✘, factor judged not suitable because either not possible to see how an adjustment could be made easily in short/medium term or relevant data unavailable

Appendix 14

Drug costs of chemotherapy regimens, with vial usage

Drug	Regimen (mg/m^2)	Dose, by BSA (mg/m^2)	Constituents parts	Cost (£)	Wastage		Last vial (% used)	Sensitivity analysis
					(mg)	(%)		
GEM (iv)	800	1360	One 1-g vial (£162.76) and two 200-mg vials (£32.55 each)	227.86	40	2.94%	80.00%	NA
	1000	1700	One 1-g vial and four 200-mg vials	292.96	100	5.88%	50.00%	NA
	1250	2125	Two 1-g vials and one 200-mg vial	358.07	75	3.53%	62.50%	NA
VNB (iv)	25	42.5	10 mg/ml: one 5-ml vial (£147.06)	147.06	7.5	17.65%	85.00%	NA
	30	51	10 mg/ml: one 5-ml vial and one 1-ml vial (£31.25)	178.31	9	17.65%	10.00%	£147.06
PAX (iv)	135	229.5	6 mg/ml: two 16.7-ml vials (£374.00 each) and one 5-ml vial (£124.79)	872.79	0.9	0.39%	97.00%	NA
	175	297.5	6 mg/ml: three 16.7-ml vials	1122.00	3.1	1.04%	96.91%	NA
	200	340	6 mg/ml: three 16.7-ml vials and two 5-ml vials	1371.58	20.6	6.06%	31.33%	NA
	250	425	6 mg/ml: four 16.7-ml vials and one 5- ml vial	1620.79	5.8	1.36%	80.67%	NA
DOC (iv)	75	127.5	40 mg/ml: one 2-ml vial (£575.00 each) and three 0.5-ml vials (£175.00 each)	1100.00	12.5	9.80%	37.50%	NA
	100	170	40 mg/ml: two 2-ml vials and one 0.5-ml vial	1325.00	10	5.88%	50.00%	NA
CDDP (iv)	60	102	Two 50-mg vials (£17.00 each) and one 10-mg vial (£4.89)	38.89	8	7.84%	20.00%	NA
	75	127.5	Two 50-mg vials and three 10-mg vials	48.67	2.5	1.96%	75.00%	NA
	80	136	Three 50-mg vials	51.00	14	10.29%	28.00%	NA
	100	170	Three 50-mg vials and two 10-mg vials	60.78	0	0.00%	100.00%	NA
	120	204	Four 50-mg vials and one 10-mg vial	72.89	6	2.94%	40.00%	NA

Notes:
1. All drug costs were taken from the BNF, September 2000
2. All calculations were based on a BSA of 1.7 m^2
3. It is assumed that, once a vial is opened, it cannot be reused and shared among patients

Appendix 15

Antiemetic regimens

Drug	Drug regimen	Antiemetic regimen	Study
GEM	GEM+BSC on days 1, 8 and 15 of 28-day cycle, for 6 cycles	Not mentioned	Anderson *et al.*, 2000[14]
	GEM on days 1, 8 and 15 of 28-day cycle, for 6 cycles	Prophylactic 5-HT_3 antagonist antiemetics not permitted at dose 1, but allowed afterward if nausea/vomiting severe. In GEM arm, 24% of patients received 5-HT_3 antagonists, 2.8% of patients received dexamethasone 20 mg, and 43% of patients received metoclopramide	Bokkel-Huinink *et al.*, 1999[24]
	GEM on days 1, 8 and 15 of 28-day cycle, for 6 cycles	Prophylactic 5-HT_3 antagonist antiemetics not permitted at dose 1, but allowed afterward if nausea/vomiting severe	Manegold *et al.*, 1997[25]
	GEM on days 1, 8 and 15 of 28-day cycle, for 6 cycles	Not mentioned	Perng *et al.*, 1997[28]
GEM+CDDP	GEM on days 1 and 8 plus CDDP on day 1 of 21-day cycle	Not mentioned	Cardenal *et al.*, 1999[26]
	GEM on days 1, 8 and 15 plus CDDP on day 2 of 28-day cycle	Along with CDDP, a programme of forced diuresis, which included at least 2 litres of fluids on day 2 of 28-day cycle, and appropriate antiemetic (5-HT_3 antagonists plus corticosteroids) and other supportive therapy	Crino *et al.*, 1999[27]
	GEM on days 1, 8 and 15 plus CDDP on day 1 of 28-day cycle	Received pretreatment iv hydration according to institutional guidelines for CDDP administration	Sandler *et al.*, 2000[29]
VNB+CDDP	VNB plus CDDP every 3 weeks	Furosemide 40 mg given if urine was delayed; metoclopramide and methyl-prednisone 120 mg recommended to prevent nausea/vomiting	Depierre *et al.*, 1994[41]
	VNB on days 1 and 8 plus CDDP on day 1 every 3–4 weeks	Not mentioned	Lorusso *et al.*, 1995[44]
	VNB weekly plus CDDP every 4 weeks	Not mentioned	Wozniak *et al.*, 1998[47]
	VNB on days 1 and 8 plus CDDP on day 1 every 3 weeks	One hour of forced diuresis after CDDP; standard antiemetic treatment (including 5-HT_3 receptor antagonists) administered before chemotherapy	Martoni *et al.*, 1998[45]
VNB	VNB weekly	Furosemide 40 mg given if urine was delayed; metoclopramide and methyl-prednisone 120 mg recommended to prevent nausea/vomiting	Depierre *et al.*, 1994[41]
	VNB weekly	Not mentioned	Lorusso *et al.*, 1995[44]
	VNB weekly	Full supportive care given	Crawford *et al.*, 1996[40]
DOC	DOC every 3 weeks	Premedicated with oral dexamethasone 8 mg twice a day for 5 days starting 24 hours before DOC therapy; antiemetics given for nausea/vomiting	Roszkowski *et al.*, 2000[22]
	DOC every 3 weeks	Premedicated with oral dexamethasone 8 mg twice a day for 5 days starting 24 hours before DOC therapy; antiemetics given for nausea/vomiting	Shepherd *et al.*, 2000[21]

continued

Drug	Drug regimen	Antiemetic regimen	Study
PAX+CDDP	PAX on day 1 plus CDDP on day 2 every 21 days	Filgrastim 5 µg/kg subcutaneously beginning on day 3 and continuing until granulocyte count was ≥ 10,000/µl	Bonomi *et al.,* 2000[30]
	PAX on day 1 plus CDDP on day 1 every 21 days	PAX preceded by oral dexamethasone 20 mg, 12 and 6 hours prior to infusion, diphenhydramine 50 mg (iv) and cimetidine 300 mg (iv) 30 minutes before infusion	Postmus *et al*., 1996[33]
	PAX on day 1 plus CDDP on day 1 every 21 days	PAX preceded by oral dexamethasone 20 mg, 12 and 6 hours prior to infusion, diphenhydramine 50 mg (iv) and cimetidine 300 mg (iv) 30 minutes before infusion; prophylactic antiemetics during and after CDDP were recommended and consisted of ondansetron and dexamethasone	Giaccone *et al*., 1998[35]
PAX	PAX every 8 hours for 24 hours, repeated every 3 weeks	Dexamethasone 20 mg orally 7 hours before and 14 hours after infusion, and diphenhydramine 50 mg and cimetidine 300 mg (iv) 1 hour before PAX	Chang *et al*., 1993[31]

Appendix 16

Details given of inpatient days/outpatient visits for chemotherapy administration in the literature

Drug	Inpatient/outpatient stay	Study	Comments
RCTs			
GEM	GEM given on days 1, 8 and 15 of each 28-day cycle, as a 30-minute infusion; dexamethasone and metoclopramide given before GEM	Perng *et al.*, 1997[28]	Outpatient clinic
	GEM (iv) over 30 minutes on days 1, 8 and 15 of each 28-day cycle, up to 6 cycles	Anderson *et al.*, 2000[14]	Outpatient clinic
	GEM (iv) over 30 minutes on days 1, 8 and 15 of each 28-day cycle, up to 6 cycles	Bokkel-Huinink *et al.*, 1999[24]	No mention of inpatient days/outpatient visits
	GEM (iv) over 30 minutes on days 1, 8 and 15 of each 28-day cycle, up to 6 cycles	Manegold *et al.*, 1997[25]	No mention of inpatient days/outpatient visits
GEM+ CDDP	GEM given on days 1 and 8, over 30 minutes, and CDDP over 60 minutes on day 1 of 21-day cycle	Cardenal *et al.*, 1999[26]	No mention of inpatient days/outpatient visits
	CDDP given on day 1 with GEM on days 1, 8 and 15 every 4 weeks	Comella *et al.*, 2000[50]	No mention of inpatient days/outpatient visits
	GEM given on days 1, 8 and 15 and CDDP on day 2 of 28-day cycle	Crino *et al.,* 1999[27]	No mention of inpatient days/outpatient visits
	GEM given on days 1, 8 and 15 and CDDP on day 1 of 28-day cycle	Sandler *et al.*, 2000[29]	No mention of inpatient days/outpatient visits
VNB	VNB infused over 20 minutes, weekly	Crawford *et al.*, 1996[40]	No mention of inpatient days/outpatient visits
	VNB infused weekly	Depierre *et al.*, 1994[41]	No mention of inpatient days/outpatient visits
	VNB infused weekly	Lorusso *et al.*, 1995[44]	No mention of inpatient days/outpatient visits
VNB+ CDDP	VNB infused over 20 minutes, CDDP infused over 60 minutes; hydration began 9 hours before treatment and continued for 12 hours afterward	Le Chevalier *et al.*, 1994[43]	No mention of inpatient days/outpatient visits
	CDDP given on days 1 and 29 and then every 6 weeks, and VNB weekly for 10 weeks	Comella *et al.*, 2000[50]	No mention of inpatient days/outpatient visits
	VNB infused weekly, with CDDP every 3 weeks	Depierre *et al.*, 1994[41]	No mention of inpatient days/outpatient visits
	CDDP given on day 1 and VNB on days 1 and 8 every 3–4 weeks	Lorusso *et al.*, 1995[44]	No mention of inpatient days/outpatient visits
	VNB (iv) given on days 1 and 8, and CDDP on day 1, every 21 days	Martoni *et al.*, 1998[45]	No mention of inpatient days/outpatient visits

continued

Drug	Inpatient/outpatient stay	Study	Comments
PAX	Dexamethasone received 12 and 6 hours before PAX, cimetidine/ranitidine (iv) 30 minutes before PAX and iphenhydramine/chloropheniramine 30 minutes before PAX; PAX (iv) given over 3 hours every 21 days	Ranson *et al.*, 2000[32]	No mention of inpatient days/outpatient visits
	PAX given every 8 hours for 24 hours (iv), every 3 weeks	Chang *et al.,* 1993[31]	No mention of inpatient days/outpatient visits
PAX+ CDDP	PAX given as 24-hour infusion on day 1, followed by 1-hour infusion of CDDP (iv) on day 2, every 21 days	Bonomi *et al.*, 2000[30]	No mention of inpatient days/outpatient visits
	CDDP given on day 1 with PAX by 3-hour infusion on day 1, every 3 weeks	Giaccone *et al.,* 1998[35]	No mention of inpatient days/outpatient visits
	CDDP given on day 1 with PAX by 3-hour infusion on day 1, every 3 weeks	Postmus *et al.*, 1996[33]	No mention of inpatient days/outpatient visits
DOC	Premedicated for 5 days (10 doses) starting 24 hours before DOC infusion, which was for 1 hour every 21 days	Shepherd *et al.*, 2000[21]	No mention of inpatient days/outpatient visits
	Premedicated for 5 days (10 doses) starting 24 hours before DOC infusion, which was for 1 hour every 21 days	Roszkowski *et al.*, 2000[22]	No mention of inpatient days/outpatient visits
Economics studies			
GEM	Outpatient stay on days 1, 8 and 15; cycle repeated every 28 days	Copley-Merriman *et al.*, 1996[57]	Costs – lab testing; infusion fees, 3 days/cycle; office visit, 3 days/cycle
	Weekly dose x 3, for 3.3 cycles	Evans, 1997[55]	Does not explicitly state that GEM patients are outpatients but assumes that this is known
	Weekly dose x 3, for 3.3 cycles	Evans, 1996[61]	Does not explicitly state that GEM patients are outpatients but assumes that this is known
	Outpatient stay on days 1, 8 and 15; cycle repeated every 28 days	Koch *et al.*, 1995[88]	Costs – physical examination, physician report, urinalysis, FBC including WBC, blood chemistry, electrolytes, prothrombin time and partial thromboplastin time, chest X-ray and ECG, transportation one-way
	Three outpatient visits	Tennvall & Fernberg, 1998[58]	
GEM+ CDDP	GEM given on days 1, 8 and 15, and CDDP on day 2 of 28-day cycle	Palmer & Brandt, 1996[59]	Assumes that patients are admitted into hospital on each day of chemotherapy treatment
VNB+ CDDP	VNB given on days 1, 8, and 15, and CDDP on day 1 of 28-day cycle	Palmer & Brandt, 1996[59]	Assumes that patients are admitted into hospital on each day of chemotherapy treatment
	VNB given weekly, and CDDP on days 1 and 29, then every 6 weeks	Hillner & Smith, 1996[63]	Patients assumed to be treated as outpatients, as was US practice in 1996
	VNB given weekly, and CDDP on days 1 and 29, then every 6 weeks	Evans, 1996[61]	High-dose CDDP administration in hospital (2 days) using inpatient schedule or VNB+CDDP cost using outpatient schedule
	VNB given weekly, and CDDP on days 1 and 29, then every 6 weeks	Evans, 1998[72]	Outpatient
PAX	PAX given as one dose every 3 weeks as 3-hour infusion during outpatient stay	Earle & Evans, 1997[89]	Outpatient
BSC		Jaakimainen *et al.*, 1990[15]	Average hospitalisation was 23.6 days per patient; based on two samples of 29 and 32 patients

Appendix 17

Hospitalisation due to chemotherapy side-effects

Drug	Regimen	Side-effect	Inpatient days	Study	Comments
GEM	GEM 1000 mg/m^2 on days 1, 8 and 15 of 28-day cycle	NA	None	Bokkel-Huinink *et al.*, 1999[24]	No hospitalisations reported for GEM arm, although hospitalisations reported in comparator (CDDP+VP-16) arm
GEM+ CDDP	No data/clinical information mentioned/available				
VNB	No data/clinical information mentioned/available				
VNB+ CDDP	No data/clinical information mentioned/available				
PAX	PAX 200 mg/m^2 every 21 days and BSC	Not specified	Not stated	Ranson *et al.*, 2000[32]	58% of patients hospitalised (PAX); perhaps proportion greater for PAX patients because they were in study for a median of 1.5 times longer than BSC patients (41% hospitalised)
PAX+ CDDP	No data/clinical information mentioned/available				
DOC	DOC 100 mg/m^2 every 3 weeks	Not specified	Not stated	Roszkowski *et al.*, 2000[22]	51% of patients hospitalised
BSC	BSC	Not specified	Not stated	Ranson *et al.*, 2000[32]	41% of patients hospitalised (BSC)
	BSC	Not specified	Not stated	Roszkowski *et al.*, 2000[22]	30% of patients hospitalised

Appendix 18

Cost per inpatient day and outpatient visit

		Cost (£)	
Cost per outpatient visit*		57	
Cost inflated to 1999/2000 prices		61	
Respiratory medicine inpatient ward costs†			
Property overheads	**Allocation method**	**Cost (£)**	**% of total cost**
Rates	Area	39,567	1%
Heating	Area	40,308	1%
Power and light	Area	12,802	0%
Building and engineering	Area	60,541	2%
Cleaning	Area	65,893	2%
Capital charges: land and buildings	Area	315,399	9%
Capital charges: equipment	Actual	51,942	2%
Total		586,452	17%
Administration overheads			
Administration: general	Pro rata to expenditure	157,106	5%
Recharge agency costs	Pro rata to expenditure	51,161	1%
Medical administration	Medical expenditure	110,286	3%
Nursing administration	Nursing expenditure	21,397	1%
Other		49,690	1%
Total		389,639	11%
Other overheads			
Laundry	Pro rata to expenditure	37,981	1%
Dietetics	Inpatient days	24,785	1%
Catering (staff)	Inpatient days	50,490	1%
Other		21,715	1%
Total		134,971	4%
Direct staff costs			
Senior medical: pay	Actual	62,543	2%
Junior medical: pay	Actual	120,713	4%
Nurses: pay	Actual	1,049,973	31%
Total		1,233,229	36%
Direct materials cost			
Drugs	Actual	183,394	
• Less cytotoxics and antiemetics		65,700	
• Residual drugs use‡		117,693	3%
Instruments and sundries	Actual	46,801	1%
Porters	Pro rata to expenditure	37,501	1%
Other		13,325	0%
Total		215,321	6%
Direct: other			
Labs costs	Actual	378,855	11%
Diagnostic radiology	Radiology use	60,165	2%
Pharmacy pay§	Residual drugs use	40,741	1%
Catering: patients	Inpatient days	92,684	3%
Other		–2,436	0%
Total		570,009	17%
Total ward costs		3,129,622	91%

continued

Other additional allocations	**Allocation method**	**Cost (£)**	**% of total cost**
Inflation	Pro rata to expenditure	202,713	6%
Health gains[‡]	Actual	102,200	3%
Superannuation	Salary costs	–22,470	–1%
Theatre Sterile Supply Unit		411	0%
Contingency	Pro rata to expenditure	17,140	0%
Total		299,994	9%
Total cost of respiratory medicine (at King's Cross Hospital)[¶]		3,429,616	100%
Number of inpatient days		25,891 days	
Cost per day		132	
Inflated to 1999/2000 prices		141	

[*] *Source: Scottish NHS Costs Blue Book, 1997. Average of 01 and 02 classification hospitals*
[†] *Courtesy of SHPIC Costing Unit*
[‡] *This item covers the cost of increases in the level of service provision for the current year*
[§] *The cost of cytotoxic drugs and the staff costs in their administration, which are costed individually elsewhere, have been removed from the ward cost calculation*
[¶] *All costs calculated for King's Cross Hospital in Dundee*

Appendix 19

Estimated costs of BSC and terminal care

	Total cost (£)	% of total cost	Number of tests/stays/ day cases/visits	Average cost per patient (£)	Average number of tests/stays/ day cases/visits per patient
BSC					
Inpatient stay	143,840	82%	1005	3,888	27
Total radiotherapy	14,573	8%		394	
Radiology scans	12,194	7%	194	330	5
Day cases	1,548	1%	11	42	0
Lab tests	1,469	1%	371	40	10
Other	1,722	1%	7	47	0
Outpatients	1,132	1%	70	31	2
Total	176,479	100%		4,470	
Number of patients: 36 Average age: 70 years					
			Terminal inpatient stay (days)	**Terminal stay cost (£)**	
Terminal care					
Average for all patients			8.06	1,341	
Median for all patients			1.00	145	
Number of patients: 36					
Cost of BSC less terminal care				3,129	
Inflated to 1999/2000 prices				3,342	

All costs courtesy of SHPIC Costing Unit

Appendix 20

BSC descriptions and caveats

Study	Description	Permitted chemotherapy?	Permitted radiotherapy?
Anderson *et al.*, 2000[14]	Any palliative treatment could be used as clinically indicated, ideally excluding chemotherapy	No?	Yes
Shepherd *et al.*, 2000[21]	Therapy was determined by treating physician	Yes	Yes
Roszkowski *et al.*, 2000[22]	As judged by treating physician. Included use of antibiotics, analgesics, transfusions or any other symptomatic treatment medically indicated. No chemotherapy or other systematic anti-cancer therapy, except radiotherapy	No	Yes
Ranson *et al.*, 2000[32]	Included palliative radiotherapy for bronchial obstruction, haemoptysis, superior vena caval obstruction and brain metastases. Corticosteroids, antibiotics, analgesics, antiemetics, transfusions and other symptomatic therapy given as required	No	Yes
Gridelli *et al.*, 1999[90]	Participating investigators were free to choose the treatment strategy	Yes	Yes

Appendix 21

Incidence of serious side-effects

	GEM	GEM+ CDDP	VNB	VNB+ CDDP	PAX	PAX+ CDDP	DOC
Blood disorders							
Neutropenia	8–16%	45–75%	4–53.2%	40–78.7%	34.0%	82–83%	28–86%
Febrile neutropenia		4.6–7%		3.0%		27–34%	1.8–22.4%
Thrombocytopenia	2–11%	20–64%	0–4%	0–30%	4.0%	5–37%	1–2%
Anaemia	4–17%	22–31%	0–1%	7–27%	3–21%	19–28%	0–16.3%
Gastrointestinal and other disorders							
Severe nausea/vomiting	3.7–37%	18–50%,	1–12.4%,	4–58%	4–5%	6–36%	1–7%
Diarrhoea	3.0%	1.0%	1–4%,	3–11.1%	8.0%	6.0%	2–3%
Alopecia	31–61%	12–13%	1–14.4%	10.7–31.9%	76.0%		
Infection	0.7–8%	4.0%			4–12.5%	8.5–14%	0–11%
References used	Anderson *et al.*, 2000[14] Bokkel-Huinink *et al.*, 1999[24] Manegold *et al.*, 1997[25] Perng *et al.*, 1997[28]	Sandler *et al.*, 2000[29] Cardenal *et al.*, 1999[26] Crino *et al.*, 1999[27] Comella *et al.*, 2000[50]	Crawford *et al.*, 1996[40] Gridelli *et al.*, 1999[90] * Lorusso *et al.*, 1995[44] Depierre *et al.*, 1994[41] Le Chevalier *et al.*, 1994[43]	Depierre *et al.*, 1994[41] Lorusso *et al.*, 1995[44] Le Chevalier *et al.*, 1994[43] Wozniak *et al.*, 1998[47] Martoni *et al.*, 1998[45] Comella *et al.*, 2000[50]	Chang *et al.*, 1993[31] Ranson *et al.*, 2000[32]	Giaccone *et al.*, 1998[35] Postmus *et al.*, 1996[33]	Shepherd *et al.*, 2000[21] Roszkowski *et al.*, 2000[22] Fossella *et al.*, 2000[23]

* *The low estimate in the Gridelli study is unusually low and may relate to the completeness of safety data in this study, which concentrated on QoL measurement*

Health Technology Assessment reports published to date

Volume 1, 1997

No. 1
Home parenteral nutrition: a systematic review.
By Richards DM, Deeks JJ, Sheldon TA, Shaffer JL.

No. 2
Diagnosis, management and screening of early localised prostate cancer.
A review by Selley S, Donovan J, Faulkner A, Coast J, Gillatt D.

No. 3
The diagnosis, management, treatment and costs of prostate cancer in England and Wales.
A review by Chamberlain J, Melia J, Moss S, Brown J.

No. 4
Screening for fragile X syndrome.
A review by Murray J, Cuckle H, Taylor G, Hewison J.

No. 5
A review of near patient testing in primary care.
By Hobbs FDR, Delaney BC, Fitzmaurice DA, Wilson S, Hyde CJ, Thorpe GH, *et al.*

No. 6
Systematic review of outpatient services for chronic pain control.
By McQuay HJ, Moore RA, Eccleston C, Morley S, de C Williams AC.

No. 7
Neonatal screening for inborn errors of metabolism: cost, yield and outcome.
A review by Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, *et al.*

No. 8
Preschool vision screening.
A review by Snowdon SK, Stewart-Brown SL.

No. 9
Implications of socio-cultural contexts for the ethics of clinical trials.
A review by Ashcroft RE, Chadwick DW, Clark SRL, Edwards RHT, Frith L, Hutton JL.

No. 10
A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment.
By Davis A, Bamford J, Wilson I, Ramkalawan T, Forshaw M, Wright S.

No. 11
Newborn screening for inborn errors of metabolism: a systematic review.
By Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, *et al.*

No. 12
Routine preoperative testing: a systematic review of the evidence.
By Munro J, Booth A, Nicholl J.

No. 13
Systematic review of the effectiveness of laxatives in the elderly.
By Petticrew M, Watt I, Sheldon T.

No. 14
When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies.
A review by Mowatt G, Bower DJ, Brebner JA, Cairns JA, Grant AM, McKee L.

Volume 2, 1998

No. 1
Antenatal screening for Down's syndrome.
A review by Wald NJ, Kennard A, Hackshaw A, McGuire A.

No. 2
Screening for ovarian cancer: a systematic review.
By Bell R, Petticrew M, Luengo S, Sheldon TA.

No. 3
Consensus development methods, and their use in clinical guideline development.
A review by Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Askham J, *et al.*

No. 4
A cost–utility analysis of interferon beta for multiple sclerosis.
By Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D.

No. 5
Effectiveness and efficiency of methods of dialysis therapy for end-stage renal disease: systematic reviews.
By MacLeod A, Grant A, Donaldson C, Khan I, Campbell M, Daly C, *et al.*

No. 6
Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model.
By Faulkner A, Kennedy LG, Baxter K, Donovan J, Wilkinson M, Bevan G.

No. 7
Antimicrobial prophylaxis in colorectal surgery: a systematic review of randomised controlled trials.
By Song F, Glenny AM.

No. 8
Bone marrow and peripheral blood stem cell transplantation for malignancy.
A review by Johnson PWM, Simnett SJ, Sweetenham JW, Morgan GJ, Stewart LA.

No. 9
Screening for speech and language delay: a systematic review of the literature.
By Law J, Boyle J, Harris F, Harkness A, Nye C.

No. 10
Resource allocation for chronic stable angina: a systematic review of effectiveness, costs and cost-effectiveness of alternative interventions.
By Sculpher MJ, Petticrew M, Kelland JL, Elliott RA, Holdright DR, Buxton MJ.

No. 11
Detection, adherence and control of hypertension for the prevention of stroke: a systematic review.
By Ebrahim S.

No. 12
Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review.
By McQuay HJ, Moore RA.

No. 13
Choosing between randomised and non-randomised studies: a systematic review.
By Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C.

No. 14
Evaluating patient-based outcome measures for use in clinical trials.
A review by Fitzpatrick R, Davey C, Buxton MJ, Jones DR.

No. 15
Ethical issues in the design and conduct of randomised controlled trials.
A review by Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J.

No. 16
Qualitative research methods in health technology assessment: a review of the literature.
By Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P.

No. 17
The costs and benefits of paramedic skills in pre-hospital trauma care.
By Nicholl J, Hughes S, Dixon S, Turner J, Yates D.

No. 18
Systematic review of endoscopic ultrasound in gastro-oesophageal cancer.
By Harris KM, Kelly S, Berry E, Hutton J, Roderick P, Cullingworth J, *et al.*

No. 19
Systematic reviews of trials and other studies.
By Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F.

No. 20
Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses.
A review by Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, *et al.*

Volume 3, 1999

No. 1
Informed decision making: an annotated bibliography and systematic review.
By Bekker H, Thornton JG, Airey CM, Connelly JB, Hewison J, Robinson MB, *et al.*

No. 2
Handling uncertainty when performing economic evaluation of healthcare interventions.
A review by Briggs AH, Gray AM.

No. 3
The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review.
By Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H.

No. 4
A randomised controlled trial of different approaches to universal antenatal HIV testing: uptake and acceptability. Annex: Antenatal HIV testing – assessment of a routine voluntary approach.
By Simpson WM, Johnstone FD, Boyd FM, Goldberg DJ, Hart GJ, Gormley SM, *et al.*

No. 5
Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review.
By Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ.

No. 6
Assessing the costs of healthcare technologies in clinical trials.
A review by Johnston K, Buxton MJ, Jones DR, Fitzpatrick R.

No. 7
Cooperatives and their primary care emergency centres: organisation and impact.
By Hallam L, Henthorne K.

No. 8
Screening for cystic fibrosis.
A review by Murray J, Cuckle H, Taylor G, Littlewood J, Hewison J.

No. 9
A review of the use of health status measures in economic evaluation.
By Brazier J, Deverill M, Green C, Harper R, Booth A.

No. 10
Methods for the analysis of quality-of-life and survival data in health technology assessment.
A review by Billingham LJ, Abrams KR, Jones DR.

No. 11
Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis.
By Zeuner D, Ades AE, Karnon J, Brown J, Dezateux C, Anionwu EN.

No. 12
Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses.
A review by Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, *et al.*

No. 13
'Early warning systems' for identifying new healthcare technologies.
By Robert G, Stevens A, Gabbay J.

No. 14
A systematic review of the role of human papillomavirus testing within a cervical screening programme.
By Cuzick J, Sasieni P, Davies P, Adams J, Normand C, Frater A, *et al.*

No. 15
Near patient testing in diabetes clinics: appraising the costs and outcomes.
By Grieve R, Beech R, Vincent J, Mazurkiewicz J.

No. 16
Positron emission tomography: establishing priorities for health technology assessment.
A review by Robert G, Milne R.

No. 17 (Pt 1)
The debridement of chronic wounds: a systematic review.
By Bradley M, Cullum N, Sheldon T.

No. 17 (Pt 2)
Systematic reviews of wound care management: (2) Dressings and topical agents used in the healing of chronic wounds.
By Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D.

No. 18
A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease.
By Berry E, Kelly S, Hutton J, Harris KM, Roderick P, Boyce JC, *et al.*

No. 19
What role for statins? A review and economic model.
By Ebrahim S, Davey Smith G, McCabe C, Payne N, Pickin M, Sheldon TA, *et al.*

No. 20
Factors that limit the quality, number and progress of randomised controlled trials.
A review by Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, *et al.*

No. 21
Antimicrobial prophylaxis in total hip replacement: a systematic review.
By Glenny AM, Song F.

No. 22
Health promoting schools and health promotion in schools: two systematic reviews.
By Lister-Sharp D, Chapman S, Stewart-Brown S, Sowden A.

No. 23
Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee.
A review by Lord J, Victor C, Littlejohns P, Ross FM, Axford JS.

Volume 4, 2000

No. 1
The estimation of marginal time preference in a UK-wide sample (TEMPUS) project.
A review by Cairns JA, van der Pol MM.

No. 2
Geriatric rehabilitation following fractures in older people: a systematic review.
By Cameron I, Crotty M, Currie C, Finnegan T, Gillespie L, Gillespie W, *et al.*

No. 3
Screening for sickle cell disease and thalassaemia: a systematic review with supplementary research.
By Davies SC, Cronin E, Gill M, Greengross P, Hickman M, Normand C.

No. 4
Community provision of hearing aids and related audiology services.
A review by Reeves DJ, Alborz A, Hickson FS, Bamford JM.

No. 5
False-negative results in screening programmes: systematic review of impact and implications.
By Petticrew MP, Sowden AJ, Lister-Sharp D, Wright K.

No. 6
Costs and benefits of community postnatal support workers: a randomised controlled trial.
By Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A.

No. 7
Implantable contraceptives (subdermal implants and hormonally impregnated intrauterine systems) versus other forms of reversible contraceptives: two systematic reviews to assess relative effectiveness, acceptability, tolerability and cost-effectiveness.
By French RS, Cowan FM, Mansour DJA, Morris S, Procter T, Hughes D, *et al.*

No. 8
An introduction to statistical methods for health technology assessment.
A review by White SJ, Ashby D, Brown PJ.

No. 9
Disease-modifying drugs for multiple sclerosis: a rapid and systematic review.
By Clegg A, Bryant J, Milne R.

No. 10
Publication and related biases.
A review by Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ.

No. 11
Cost and outcome implications of the organisation of vascular services.
By Michaels J, Brazier J, Palfreyman S, Shackley P, Slack R.

No. 12
Monitoring blood glucose control in diabetes mellitus: a systematic review.
By Coster S, Gulliford MC, Seed PT, Powrie JK, Swaminathan R.

No. 13
The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature.
By Elkan R, Kendrick D, Hewitt M, Robinson JJA, Tolley K, Blair M, *et al.*

No. 14
The determinants of screening uptake and interventions for increasing uptake: a systematic review.
By Jepson R, Clegg A, Forbes C, Lewis R, Sowden A, Kleijnen J.

No. 15
The effectiveness and cost-effectiveness of prophylactic removal of wisdom teeth.
A rapid review by Song F, O'Meara S, Wilson P, Golder S, Kleijnen J.

No. 16
Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views.
By Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, *et al.*

No. 17
A rapid and systematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast and ovarian cancer.
By Lister-Sharp D, McDonagh MS, Khan KS, Kleijnen J.

No. 18
Liquid-based cytology in cervical screening: a rapid and systematic review.
By Payne N, Chilcott J, McGoogan E.

No. 19
Randomised controlled trial of non-directive counselling, cognitive–behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care.
By King M, Sibbald B, Ward E, Bower P, Lloyd M, Gabbay M, *et al.*

No. 20
Routine referral for radiography of patients presenting with low back pain: is patients' outcome influenced by GPs' referral for plain radiography?
By Kerry S, Hilton S, Patel S, Dundas D, Rink E, Lord J.

No. 21
Systematic reviews of wound care management: (3) antimicrobial agents for chronic wounds; (4) diabetic foot ulceration.
By O'Meara S, Cullum N, Majid M, Sheldon T.

No. 22
Using routine data to complement and enhance the results of randomised controlled trials.
By Lewsey JD, Leyland AH, Murray GD, Boddy FA.

No. 23
Coronary artery stents in the treatment of ischaemic heart disease: a rapid and systematic review.
By Meads C, Cummins C, Jolly K, Stevens A, Burls A, Hyde C.

No. 24
Outcome measures for adult critical care: a systematic review.
By Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, *et al.*

No. 25
A systematic review to evaluate the effectiveness of interventions to promote the initiation of breastfeeding.
By Fairbank L, O'Meara S, Renfrew MJ, Woolridge M, Sowden AJ, Lister-Sharp D.

No. 26
Implantable cardioverter defibrillators: arrhythmias. A rapid and systematic review.
By Parkes J, Bryant J, Milne R.

No. 27
Treatments for fatigue in multiple sclerosis: a rapid and systematic review.
By Brañas P, Jordan R, Fry-Smith A, Burls A, Hyde C.

No. 28
Early asthma prophylaxis, natural history, skeletal development and economy (EASE): a pilot randomised controlled trial.
By Baxter-Jones ADG, Helms PJ, Russell G, Grant A, Ross S, Cairns JA, *et al.*

No. 29
Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis.
By Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW.

No. 30
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists in the medical management of unstable angina.
By McDonagh MS, Bachmann LM, Golder S, Kleijnen J, ter Riet G.

No. 31
A randomised controlled trial of prehospital intravenous fluid replacement therapy in serious trauma.
By Turner J, Nicholl J, Webber L, Cox H, Dixon S, Yates D.

No. 32
Intrathecal pumps for giving opioids in chronic pain: a systematic review.
By Williams JE, Louw G, Towlerton G.

No. 33
Combination therapy (interferon alfa and ribavirin) in the treatment of chronic hepatitis C: a rapid and systematic review.
By Shepherd J, Waugh N, Hewitson P.

No. 34
A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies.
By MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AMS.

No. 35
Intravascular ultrasound-guided interventions in coronary artery disease: a systematic literature review, with decision-analytic modelling, of outcomes and cost-effectiveness.
By Berry E, Kelly S, Hutton J, Lindsay HSJ, Blaxill JM, Evans JA, *et al.*

No. 36
A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression.
By Simpson S, Corney R, Fitzgerald P, Beecham J.

No. 37
Systematic review of treatments for atopic eczema.
By Hoare C, Li Wan Po A, Williams H.

No. 38
Bayesian methods in health technology assessment: a review.
By Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR.

No. 39
The management of dyspepsia: a systematic review.
By Delaney B, Moayyedi P, Deeks J, Innes M, Soo S, Barton P, *et al.*

No. 40
A systematic review of treatments for severe psoriasis.
By Griffiths CEM, Clark CM, Chalmers RJG, Li Wan Po A, Williams HC.

Volume 5, 2001

No. 1
Clinical and cost-effectiveness of donepezil, rivastigmine and galantamine for Alzheimer's disease: a rapid and systematic review.
By Clegg A, Bryant J, Nicholson T, McIntyre L, De Broe S, Gerard K, *et al.*

No. 2
The clinical effectiveness and cost-effectiveness of riluzole for motor neurone disease: a rapid and systematic review.
By Stewart A, Sandercock J, Bryan S, Hyde C, Barton PM, Fry-Smith A, *et al.*

No. 3
Equity and the economic evaluation of healthcare.
By Sassi F, Archard L, Le Grand J.

No. 4
Quality-of-life measures in chronic diseases of childhood.
By Eiser C, Morse R.

No. 5
Eliciting public preferences for healthcare: a systematic review of techniques.
By Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM, *et al.*

No. 6
General health status measures for people with cognitive impairment: learning disability and acquired brain injury.
By Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J.

No. 7
An assessment of screening strategies for fragile X syndrome in the UK.
By Pembrey ME, Barnicoat AJ, Carmichael B, Bobrow M, Turner G.

No. 8
Issues in methodological research: perspectives from researchers and commissioners.
By Lilford RJ, Richardson A, Stevens A, Fitzpatrick R, Edwards S, Rock F, *et al.*

No. 9
Systematic reviews of wound care management: (5) beds; (6) compression; (7) laser therapy, therapeutic ultrasound, electrotherapy and electromagnetic therapy.
By Cullum N, Nelson EA, Flemming K, Sheldon T.

No. 10
Effects of educational and psychosocial interventions for adolescents with diabetes mellitus: a systematic review.
By Hampson SE, Skinner TC, Hart J, Storey L, Gage H, Foxcroft D, *et al.*

No. 11
Effectiveness of autologous chondrocyte transplantation for hyaline cartilage defects in knees: a rapid and systematic review.
By Jobanputra P, Parry D, Fry-Smith A, Burls A.

No. 12
Statistical assessment of the learning curves of health technologies.
By Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT.

No. 13
The effectiveness and cost-effectiveness of temozolomide for the treatment of recurrent malignant glioma: a rapid and systematic review.
By Dinnes J, Cave C, Huang S, Major K, Milne R.

No. 14
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of debriding agents in treating surgical wounds healing by secondary intention.
By Lewis R, Whiting P, ter Riet G, O'Meara S, Glanville J.

No. 15
Home treatment for mental health problems: a systematic review.
By Burns T, Knapp M, Catty J, Healey A, Henderson J, Watt H, *et al.*

No. 16
How to develop cost-conscious guidelines.
By Eccles M, Mason J.

No. 17
The role of specialist nurses in multiple sclerosis: a rapid and systematic review.
By De Broe S, Christopher F, Waugh N.

No. 18
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of orlistat in the management of obesity.
By O'Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 19
The clinical effectiveness and cost-effectiveness of pioglitazone for type 2 diabetes mellitus: a rapid and systematic review.
By Chilcott J, Wight J, Lloyd Jones M, Tappenden P.

No. 20
Extended scope of nursing practice: a multicentre randomised controlled trial of appropriately trained nurses and preregistration house officers in pre-operative assessment in elective general surgery.
By Kinley H, Czoski-Murray C, George S, McCabe C, Primrose J, Reilly C, *et al.*

No. 21
Systematic reviews of the effectiveness of day care for people with severe mental disorders: (1) Acute day hospital versus admission; (2) Vocational rehabilitation; (3) Day hospital versus outpatient care.
By Marshall M, Crowther R, Almaraz-Serrano A, Creed F, Sledge W, Kluiter H, *et al.*

No. 22
The measurement and monitoring of surgical adverse events.
By Bruce J, Russell EM, Mollison J, Krukowski ZH.

No. 23
Action research: a systematic review and guidance for assessment.
By Waterman H, Tillen D, Dickson R, de Koning K.

No. 24
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of gemcitabine for the treatment of pancreatic cancer.
By Ward S, Morris E, Bansback N, Calvert N, Crellin A, Forman D, *et al.*

No. 25
A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer.
By Lloyd Jones M, Hummel S, Bansback N, Orr B, Seymour M.

No. 26
Comparison of the effectiveness of inhaler devices in asthma and chronic obstructive airways disease: a systematic review of the literature.
By Brocklebank D, Ram F, Wright J, Barry P, Cates C, Davies L, *et al.*

No. 27
The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint.
By Bryan S, Weatherburn G, Bungay H, Hatrick C, Salas C, Parry D, *et al.*

No. 28
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of topotecan for ovarian cancer.
By Forbes C, Shirran L, Bagnall A-M, Duffy S, ter Riet G.

No. 29
Superseded by a report published in volume 6.

No. 30
The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial.
By Kendrick D, Fielding K, Bentley E, Miller P, Kerslake R, Pringle M.

No. 31
Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients.
By McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, *et al.*

No. 32
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer.
By Clegg A, Scott DA, Sidhu M, Hewitson P, Waugh N.

Health Technology Assessment Programme

Prioritisation Strategy Group

Members

Chair
Professor Kent Woods
Director,
NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Professor Bruce Campbell
Consultant General Surgeon
Royal Devon & Exeter Hospital

Professor Shah Ebrahim
Professor of Epidemiology
of Ageing
University of Bristol

Dr John Reynolds
Clinical Director
Acute General Medicine SDU
Oxford Radcliffe Hospital

Dr Ron Zimmern
Director, Public Health
Genetics Unit
Strangeways Research
Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director
Professor Kent Woods
Director, NHS HTA
Programme, &
Professor of Therapeutics
University of Leicester

Chair
Professor Shah Ebrahim
Professor of Epidemiology
of Ageing
University of Bristol

Deputy Chair
Professor Jon Nicholl
Director, Medical Care
Research Unit
University of Sheffield

Professor Douglas Altman
Director, ICRF Medical
Statistics Group
University of Oxford

Professor John Bond
Director, Centre for Health
Services Research
University of Newcastle-
upon-Tyne

Ms Christine Clark
Freelance Medical Writer
Bury, Lancs

Professor Martin Eccles
Professor of
Clinical Effectiveness
University of Newcastle-
upon-Tyne

Dr Andrew Farmer
General Practitioner &
NHS R&D
Clinical Scientist
Institute of Health Sciences
University of Oxford

Professor Adrian Grant
Director, Health Services
Research Unit
University of Aberdeen

Dr Alastair Gray
Director, Health Economics
Research Centre
Institute of Health Sciences
University of Oxford

Professor Mark Haggard
Director, MRC Institute
of Hearing Research
University of Nottingham

Professor Jenny Hewison
Senior Lecturer
School of Psychology
University of Leeds

Professor Alison Kitson
Director, Royal College of
Nursing Institute, London

Dr Donna Lamping
Head, Health Services
Research Unit
London School of Hygiene
& Tropical Medicine

Professor David Neal
Professor of Surgery
University of Newcastle-
upon-Tyne

Professor Gillian Parker
Nuffield Professor of
Community Care
University of Leicester

Dr Tim Peters
Reader in Medical Statistics
University of Bristol

Professor Martin Severs
Professor in Elderly
Health Care
University of Portsmouth

Dr Sarah Stewart-Brown
Director, Health Services
Research Unit
University of Oxford

Professor Ala Szczepura
Director, Centre for Health
Services Studies
University of Warwick

Dr Gillian Vivian
Consultant in Nuclear
Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Professor Graham Watt
Department of
General Practice
University of Glasgow

Dr Jeremy Wyatt
Senior Fellow
Health Knowledge
Management Centre
University College London

continued

continued

Diagnostic Technologies & Screening Panel

Members

Chair
Dr Ron Zimmern
Director, Public Health Genetics Unit
Strangeways Research Laboratories
Cambridge

Dr Philip J Ayres
Consultant in Epidemiology & Public Health
The Leeds Teaching Hospitals NHS Trust

Mrs Stella Burnside
Chief Executive, Altnagelvin Hospitals Health & Social Services Trust
Londonderry
Northern Ireland

Dr Paul O Collinson
Consultant Chemical Pathologist & Senior Lecturer
St George's Hospital, London

Dr Barry Cookson
Director, Laboratory of Hospital Infection
Public Health Laboratory Service, London

Professor Howard Cuckle
Professor of Reproductive Epidemiology
University of Leeds

Dr Carol Dezateux
Senior Lecturer in Paediatric Epidemiology
Institute of Child Health
London

Professor Adrian K Dixon
Professor of Radiology
Addenbrooke's Hospital
Cambridge

Mr Steve Ebdon-Jackson
Head, Diagnostic Imaging & Radiation Protection Team
Department of Health, London

Dr Tom Fahey
Senior Lecturer in General Practice
University of Bristol

Dr Andrew Farmer
General Practitioner & NHS Clinical Scientist
Institute of Health Sciences
University of Oxford

Mrs Gillian Fletcher
Antenatal Teacher & Tutor
National Childbirth Trust
Reigate

Professor Jane Franklyn
Professor of Medicine
University of Birmingham

Dr JA Muir Gray
Joint Director, National Screening Committee
NHS Executive, Oxford

Dr Peter Howlett
Executive Director – Development
Portsmouth Hospitals NHS Trust

Professor Alistair McGuire
Professor of Health Economics
City University, London

Mrs Kathlyn Slack
Professional Support
Diagnostic Imaging & Radiation Protection Team
Department of Health
London

Mr Tony Tester
Chief Officer, South Bedfordshire Community Health Council
Luton

Pharmaceuticals Panel

Members

Chair
Dr John Reynolds
Clinical Director – Acute General Medicine SDU
Oxford Radcliffe Hospital

Dr Felicity J Gabbay
Managing Director, Transcrip Ltd
Milford-on-Sea, Hants

Mr Peter Golightly
Director, Trent Drug Information Services
Leicester Royal Infirmary

Dr Alastair Gray
Director, Health Economics Research Centre
Institute of Health Sciences
University of Oxford

Mrs Jeannette Howe
Senior Principal Pharmacist
Department of Health, London

Dr Andrew Mortimore
Consultant in Public Health Medicine
Southampton & South West Hants Health Authority

Mr Nigel Offen
Head of Clinical Quality
NHS Executive – Eastern
Milton Keynes

Professor Robert Peveler
Professor of Liaison Psychiatry
Royal South Hants Hospital
Southampton

Mrs Marianne Rigge
Director, College of Health
London

Dr Frances Rotblat
Manager, Biotechnology Group
Medicines Control Agency
London

Mr Bill Sang
Chief Executive
Salford Royal Hospitals NHS Trust

Dr Eamonn Sheridan
Consultant in Clinical Genetics
St James's University Hospital
Leeds

Mrs Katrina Simister
New Products Manager
National Prescribing Centre
Liverpool

Dr Ross Taylor
Senior Lecturer
Department of General Practice & Primary Care
University of Aberdeen

Dr Richard Tiner
Medical Director
Association of the British Pharmaceutical Industry
London

Professor Jenifer Wilson-Barnett
Head, Florence Nightingale Division of Nursing & Midwifery
King's College, London

Mr David J Wright
Chief Executive
International Glaucoma Association, London

Therapeutic Procedures Panel

Members

Chair
Professor Bruce Campbell
Consultant General Surgeon
Royal Devon & Exeter Hospital

Professor John Bond
Professor of
Health Services Research
University of Newcastle-upon-Tyne

Ms Judith Brodie
Head of Cancer
Support Service
Cancer BACUP, London

Ms Tracy Bury
Head of Research
& Development
Chartered Society of
Physiotherapy, London

Mr Michael Clancy
Consultant in A&E Medicine
Southampton General Hospital

Professor Collette Clifford
Professor of Nursing
University of Birmingham

Dr Katherine Darton
Information Unit
MIND – The Mental Health
Charity, London

Mr John Dunning
Consultant Cardiothoracic
Surgeon
Papworth Hospital NHS Trust
Cambridge

Mr Jonothan Earnshaw
Consultant Vascular Surgeon
Gloucestershire Royal Hospital

Professor David Field
Professor of Neonatal Medicine
The Leicester Royal Infirmary
NHS Trust

Professor FD Richard Hobbs
Professor of Primary Care
& General Practice
University of Birmingham

Mr Richard Johanson
Consultant & Senior Lecturer
North Staffordshire Infirmary
NHS Trust, Stoke-on-Trent

Dr Duncan Keeley
General Practitioner
Thame, Oxon

Dr Phillip Leech
Principal Medical Officer
Department of Health, London

Professor James Lindesay
Professor of Psychiatry for
the Elderly
University of Leicester

Professor Rajan Madhok
Director of Health Policy
& Public Health
East Riding & Hull
Health Authority

Dr Mike McGovern
Branch Head
Department of Health
London

Dr John C Pounsford
Consultant Physician
Frenchay Healthcare Trust
Bristol

Dr Mark Sculpher
Senior Research Fellow in
Health Economics
University of York

Dr Ken Stein
Consultant in Public
Health Medicine
North & East Devon
Health Authority, Exeter

Expert Advisory Network

Members

Professor John Brazier
Director of Health Economics
University of Sheffield

Mr Shaun Brogan
Chief Executive, Ridgeway
Primary Care Group
Aylesbury, Bucks

Mr John A Cairns
Director, Health Economics
Research Unit
University of Aberdeen

Dr Nicky Cullum
Reader in Health Studies
University of York

Professor Pam Enderby
Chair of Community
Rehabilitation
University of Sheffield

Mr Leonard R Fenwick
Chief Executive
Freeman Hospital
Newcastle-upon-Tyne

Ms Grace Gibbs
Deputy Chief Executive
West Middlesex
University Hospital

Dr Neville Goodman
Consultant Anaesthetist
Southmead Hospital, Bristol

Professor Robert E Hawkins
CRC Professor & Director of
Medical Oncology
Christie Hospital NHS Trust
Manchester

Professor Allen Hutchinson
Director of Public Health &
Deputy Dean, ScHARR
University of Sheffield

Professor David Mant
Professor of General Practice
Institute of Health Sciences
University of Oxford

Professor Alexander Markham
Director
Molecular Medicine Unit
St James's University Hospital
Leeds

Dr Chris McCall
General Practitioner
Corfe Mullen, Dorset

Dr Peter Moore
Freelance Science Writer
Ashtead, Surrey

Dr Sue Moss
Associate Director, Cancer
Screening Evaluation Unit
Institute of Cancer Research
Sutton, Surrey

Mrs Julietta Patnick
National Coordinator
NHS Cancer Screening
Programmes, Sheffield

Professor Jennie Popay
Professor of Sociology &
Community Health
University of Salford

Professor Chris Price
Professor of
Clinical Biochemistry
St Bartholomew's & The
Royal London School of
Medicine & Dentistry

Mr Simon Robbins
Chief Executive
Camden & Islington
Health Authority, London

Dr William Rosenberg
Senior Lecturer &
Consultant in Medicine
University of Southampton

Dr Sarah Stewart-Brown
Director, Health Services
Research Unit
University of Oxford

Dr Gillian Vivian
Consultant in Nuclear
Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Mrs Joan Webster
Former Chair
Southern Derbyshire
Community Health Council
Nottingham

Feedback

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278