
1 Sahar Mokhles  
2 Fergus Macbeth  
3 Tom Treasure*  
4 Riad N Younes  
5 Robert C Rintoul  
6 Francesca Fiorentino†  
7 Ad J.J.C. Bogers  
8 Johanna J. M. Takkenberg

1 Cardio-thoracic Surgery, Erasmus MC Rotterdam  
2 Wales Cancer Trials Unit, Cardiff University, Cardiff, UK  
3 Clinical Operational Research Unit, University College London  
4 Hospital Alemão Oswaldo Cruz, São Paulo, Brazil  
5 Thoracic Oncology, Papworth Hospital, Cambridge, UK  
6 Imperial College Trials Unit & Division of Surgery, Imperial College London, London, UK

†Dr Fiorentino is part funded by the British Heart Foundation  
*Corresponding Author

Professor Tom Treasure  
Clinical Operational Research Unit UCL  
4 Taviton Street WC1H 0BT  
London UK

Phone/fax 01233 740 378  
E-mail tom.treasure@gmail.com

Key words: Lung cancer, surgery, lymph node staging

Competing Interest Statement: the authors have none.

Word count 2996 text of 4743 words in total manuscript
Structured Abstract (234/250 words)

Objectives
To re-examine the evidence for recommendations for complete dissection versus sampling of ipsilateral mediastinal lymph nodes during lobectomy for cancer.

Methods
We searched for randomised trials of systematic mediastinal lymphadenectomy versus mediastinal sampling. We performed a textual analysis of the authors’ own starting assumptions and conclusion. We analysed the trial designs and risk of bias. We extracted data on early mortality, perioperative complications, overall survival, local recurrence and distant recurrence for meta-analysis.

Results
We found five randomised controlled trials recruiting 1,980 patients spanning 1989 to 2007. The expressed starting position in 3/5 studies was a conviction that systematic dissection was effective. Long-term survival was better with lymphadenectomy compared with sampling (Hazard Ratio 0.78; 95% CI 0.69-0.89) as was perioperative survival (Odds Ratio 0.59; 95% CI 0.25-1.36, non-significant). But there was an overall high risk of bias and a lack of intention to treat analysis. There were higher rates (non-significant) of perioperative complications including bleeding, chylothorax and recurrent nerve palsy with lymphadenectomy.

Conclusions
The high risk of bias in these trials makes the overall conclusion insecure. The finding of clinically important surgically related morbidities but lower perioperative mortality with lymphadenectomy seems inconsistent. The multiple variables in patients, cancers and available treatments suggest that large pragmatic multicentre trials, testing currently available strategies, are the best way to find out which are more effective. The number of patients affected with lung cancer makes trials feasible.
Introduction

The surgical approach to ipsilateral mediastinal (N2) nodes at the time of lobectomy for lung cancer has long been a subject of interest. The European Society of Thoracic Surgeons (ESTS) Guidelines in 2006 stated “adherence to these guidelines will standardize the intraoperative lymph node staging and pathologic evaluation, and improve pathologic staging, which will help decide on the best adjuvant therapy”. [1] The opening statement of the International Association for the Study of Lung Cancer (IASLC) staging project’s Proposals for the Revision of the N Descriptors in the 8th Edition of the Tumor Node Metastasis (TNM) Classification for Lung Cancer reads: ‘Nodal status is considered to be one of the most reliable indicators of the prognosis in patients with lung cancer and thus is indispensable in determining the optimal therapeutic options.’[2] The extent of nodal dissection and the number of nodes removed and sent to the pathology laboratory is used as a quality standard in some jurisdictions.

Arguments in favour of more extensive lymph nodes dissection fall into three groups.

1. More accurate N staging makes research comparisons between treatment effects more reliable.
2. More complete N staging provides more information on which to plan already available and novel adjuvant treatments.
3. Removal of unsuspected or microscopic cancer by complete lymphadenectomy maximises the possibility of cure.

There can be little doubt that systematic ipsilateral mediastinal lymphadenectomy, rather than lymph node sampling protocols, maximises the information available for pathological staging as far as the ipsilateral mediastinum is concerned. However in the era of modern imaging and less invasive biopsies, how much it actually adds to staging is open to question.[3;4] Furthermore, an operation for lung resection through either thoracotomy or videothoracoscopy, offers no opportunity to sample nodes on the other side of the chest. These can and, if necessary, should be assessed preoperatively by imaging and one or more of the minimally invasive biopsy techniques now available.
The argument that the chance of additional cures by removal of otherwise undetected lymph node metastases has prompted recent discussion. Lim and eminent European colleagues have argued cogently that if low volume N2 disease does not preclude lung resection then mediastinal dissection at the time of thoracotomy spares the patient preoperative biopsies.[5] There appear to be substantial transatlantic differences as outlined by Rocco and colleagues: “North American surgeons are more likely to surgically stage the mediastinum before operation, are less likely to offer surgical treatment when N2 disease is identified preoperatively, and are more likely to use induction therapy before resection. By contrast, European surgeons may offer operation as the initial treatment followed by adjuvant therapy in selected cases of N2 disease, and they may perform a more aggressive intraoperative nodal dissection.”[6]

Furthermore with pressure to reduce the burden of surgery in frail elderly patients or in the presence of comorbidities there is increasing interest in treatment with stereotactic ablative radiotherapy (SABR/SBRT).[7] Full pathological N2 staging is not possible, at least not as part of the therapeutic intervention, making it not equivalent to surgery. The same argument has been raised against videothoracoscopy (VATS) but has largely been resolved by evidence that surgeons experienced in VATS can achieve the required nodal clearance standards. [8;9] If mediastinal dissection is used as a reason for not moving to less invasive means of treating lung cancer, this should be based on sound evidence in the interests of patients.

The use of protocols for mediastinal lymph node dissection (MLND) and mediastinal lymph node sampling (MNLS) have been studied in randomised controlled trials. Four RCTs[10-13] were included in a meta-analysis reported in late 2014.[14] The authors concluded “Results for overall survival, local recurrence rate, and distant metastasis rate were similar between MLND and MLNS in early stage NSCLC patients. There was no evidence that MLND increased complications compared with MLNS. Whether or not MLND is superior to MLNS for stage II–IIIA remains to be determined.” We have added a fifth study[15] and performed a detailed analysis of the text and the data.
Materials and Methods:

Search strategy and selection of studies

A systematic review of literature on surgical policy with respect to mediastinal lymph node sampling or radical lymph node dissection in patients with primary lung cancer was conducted according to the PRISMA guidelines.[16;17] This selection of studies for inclusion was based on predefined eligibility criteria and conducted according to a predefined methodological approach.

Search strategy

An extensive search for published articles was conducted on May 1st 2015 in collaboration with a medical librarian, using among others the electronic databases Medline (Ovid), Embase.com, the Cochrane library and Web of Science. A total of ten databases were searched from inception until May 2015 and updated in April 2016. The main search terms were chosen to identify ‘non-small cell lung cancer’ and ‘mediastinal lymph node dissection or sampling’. Appropriate thesaurus terms (for Medline, Embase and CINAHL) and words and phrases in title and/or abstract were combined by Boolean logical operators and adapted to the appropriate syntax of each databases. (Full details of databases used, and the syntax for each database, are available as Appendix 1).

Selection of studies

The resulting papers were then screened manually for relevance by two independent investigators (SM and TT). Any disagreement about including a paper, was to be resolved by discussion with RY. Studies were included if they reported comparisons of randomly assigned groups of patients undergoing mediastinal lymph node dissection or sampling for non-small cell lung cancer. We limited our search to studies that were conducted in humans, published in the last 35 years and written in English. We excluded studies not providing analysable data on survival. To ensure that no potentially valid studies were missed, the reference lists of relevant reviews and included studies were cross-checked.
Data extraction

Data were extracted by two of the investigators (SM and TT) using standardised tables developed for this purpose and independently checked by another investigator (RY). From each study we collected the number of patients, patient baseline characteristics, recurrence rates, and overall survival. The risk of bias was assessed (by SM and FM) using the Cochrane Handbook [18] and from information available in the publications. The authors’ prior position, the vulnerability of the study design to bias, and the authors’ own interpretation of their results were extracted from the text.

Statistical analysis

Overall survival data were extracted as event rates following systematic mediastinal lymph node dissection versus mediastinal lymph node sampling of all randomised comparisons. Where possible hazard ratios (HR) were derived from Kaplan-Meier curves. The method described by Williamson et al [19] was used to estimate a logarithmic HR with corresponding variance when the number of patients at risk was given at each time frame. If these data were not provided, the method described by Parmar et al [20] was used. For each study, we used a spreadsheet programmed to estimate the overall HR with 95% confidence intervals (CI) using an inverse variance-weighted average.[21] Whereas OR was derived from the percentages of deaths in each arm at the time of reporting (early mortality), the HR gives an estimate of the overall relative survival which is more relevant when considering a time to event endpoint. HR was used to calculate absolute mortality risk reduction at 5 years. To illustrate early mortality and complications we used OR as these outcomes are not time-to-event outcomes and therefore differences in length of follow up, the number and timing of events does not have to be taken into account.[21]

Reported study characteristics are presented as numbers or percentages in tables. The linearized occurrence rate (LOR) for each late mortality was calculated by dividing the number of deaths by the total follow-up time in patient-years, and then pooled on a logarithmic scale using the inverse variance method within a random-effects model. The pooled LOR was used to estimate the absolute mortality risk reduction at 5 years. Heterogeneity among the included studies was analysed with the $I^2$
measure with values of 25%, 50%, and 75% taken to represent, respectively, low, moderate, and high heterogeneity.[18]. Statistical analyses were performed using Review Manager for Windows.[22]
Results

Figure 1 illustrates the literature search process. After removal of duplicates, 2489 titles and abstracts were screened. After successive exclusions there were nine papers [10-13;15;23-26] reporting five randomised trials from which data were extracted for meta-analysis.

Technical definitions of the procedures in all included studies are provide in Appendix 2 and surgical procedures in Appendix 3.

There are variations in the words used and hence in the abbreviations. In the authors’ abbreviations S variably stands for either ‘sampling’ or ‘systematic’ which are opposites in the context of this analysis. The essential difference under test is between systematic mediastinal lymph node dissection to achieve complete lymphadenectomy, identified in our analysis as [MLND] and lymph node sampling abbreviated to [MLNS]. D for dissection, when used, signifies a systematic lymphadenectomy.

In Table 1 we have extracted from the text an indication of the authors’ prior position and a summary of their own conclusions.

Risk of bias

Table 2 shows that all five trials were at risk of bias with the methods for sequence generation and allocation concealment unclear in all and a failure to carry out an intention to treat analysis in three.

Results of the meta-analysis

For perioperative survival (Fig.2a) there was an overall non-significant difference in favour of the more radical arms [MLND] compared with sampling [MLNS] (Odds Ratio for death 0.59 (95% CI 0.25-1.36)). This was largely due to the ACOSOG Z0031 trial.

Overall survival (Fig.2b) was greater after mediastinal dissection than after sampling (HR 0.78 (95% CI 0.69-0.89) Absolute mortality risk reduction at 5 years was calculated using the LOR calculated
from the HR. For the [MLND] group the pooled LOR was 0.0688 (i.e. late mortality of 6.88% per year) and for the [MLNS] group this was 0.578 (i.e. late mortality of 5.78% per year). We have considered these LOR from three studies in the MLND and MLNS groups as the most reliable estimate of late mortality.[10-12] Absolute mortality risk at 5 years for the MLNS group was 34.4%. A HR of 0.78 (Fig.2b) was considered as the baseline risk for overall mortality, and this information was used to calculated the relative mortality risk reduction (MLND compared to MLNS) of 0.22. The relative mortality risk reduction and 5 year risk of death in the MLNS group resulted in absolute mortality risk reduction of 7.6% in favour of MLND group.

Local recurrence (Fig.2c) was non-significantly lower after MLND (55/900; 6.1%) than sampling (75/878; 8.5%. P=0.12). Distant recurrence (Fig.2d) was also non-significantly lower after MLND (191/900; 21.2%) rather than sampling (219/878; 24.9%. P=0.07).

However complications (Fig.3) were generally higher after dissection than after sampling. Bleeding 4% versus 2.8%; bronchial secretions 12.1% versus 7.7%; chylothorax 1.8% versus 0.7%; recurrent laryngeal nerve injury 2.4% versus 1.1%. As expected, the burden of complications (Fig.3) is greater for MLND due to the more extensive dissection. These included bleeding, chylothorax and recurrent nerve injury.
The main objective of additional, more complex surgery is to provide a benefit that outweighs any additional risk. In this meta-analysis of 1,980 patients the hazard ratio for overall survival was 0.78 (95% CI 0.69 to 0.89) favouring systematic lymphadenectomy [MLND] rather than sampling [MLNS] and this equates with an absolute reduction in risk of death at 5 years of 7.6%. (Fig.2b) If these data are reliable this would be clinically significant confirming this procedure as standard. It would also provide a caveat about equivalence of SABR/SBRT instead of surgery for primary lung cancer. There are however a number of things that reduce confidence in the validity of this conclusion.

How do we explain the better perioperative survival (Fig.2a) associated with the more extensive lymphadenectomy [MLND]? This is counterintuitive and is made more so by the tally of complications. (Fig.3) As might be expected, bleeding (P=0.36), chylothorax (P=0.08) and recurrent nerve injury (P=0.14) were all more frequent with the more extensive surgery; although not statistically significant in this analysis they are anticipated complications of more extensive surgery in the mediastinum. Despite the excess morbidity with [MLND] the early mortality was lower. In unblinded trials, run by doctors with a vested interest in the outcome, there are opportunities for reassignment or exclusion of patients in trials. The exercise of bias may be unintentional but later we will discuss data which suggest it may have happened.

These five trials were intended to test in survival terms the effectiveness of extending the surgery performed at the time of lobectomy to include lymphadenectomy. This has direct bearing on three distinct drives for change in clinical practice.

1. When stereotactic radiotherapy is used as treatment for primary lung cancer rather than lobectomy[27] lymphadenectomy is precluded.

2. When videothoracoscopic surgery is used instead of open lobectomy, the prior assumption is that lymphadenectomy is less often complete.[8]
3. An increasing role of lymphadenectomy will be to provide more tissue and more complete staging to guide multimodality therapy.[28]

Despite a difference in overall survival, lymphadenectomy was not associated with a significant reduction in the rates of either local or distant recurrence and we cannot infer from the trials whether the apparent effect on survival is due to removal of more involved nodes having a beneficial effect on survival or the information from more accurate nodal staging guiding adjuvant treatment with consequent benefit. Only three studies mention the use of post-operative radiotherapy and it is not clear if the rates of use varied. Chemotherapy is not mentioned in any of the reports of three of the trials.[11;13;15;23;24] Use of preoperative chemotherapy was an exclusion criterion in one of the trials[26] and was used in a few cases where small-cell lung cancer or a non-lung primary was the cause of mediastinal nodal metastases.[12] It is not clear whether or not adjuvant chemotherapy was given to patients with N2 disease in any of the studies; this might have made a different in outcomes.

It is also possible that the additional knowledge concerning staging obtained during the study influenced the composition of the reported trial arms in two of the studies. In the ACOSOG Z0030 trial all patients had sampling and frozen section and the protocol required patients with any positive nodes to not be randomised.[26] We are not told how many patients were excluded in this process and we cannot estimate what effect, if any that would have on the conclusions. After randomisation and presumably in the knowledge of findings during the trial “retrospective review found 155 patients to be ineligible for participation”. It appears that this was a decision which included knowledge of pTNM thus nullifying the intention to treat principle. This revision of the assigned arms took out 14% of randomised patients (155/1111) and overall there was an imbalance of 5% between the arms.

In the table of staging provided in the report by Wu and colleagues [13] the distribution between stages I, II and III was 42%, 30% and 28% for patients having sampling but was 24%, 28% and 48%
for patients having systemic nodal dissection. In the design of the trial these should have been
generated to clinical staging (cTNM). We suspect that the intraoperative findings may have been
used to restage the patients by pTNM thus inadvertently violating the randomisation process by
reassigning the patients on the basis of trial findings. The revised staging has subsequently been used
to make stage specific comparisons which are therefore erroneous.[13] If there is a 20% stage shift
between the three stages, occult N2 disease, undiscovered by sampling is very common. What we
cannot deduce is whether mediastinal nodal dissection will then alter the outcome for the patient. This
illustrates the distinction to be made between ‘efficacy’ and ‘effectiveness’ as used in evidence based
medicine. The efficacy of removing more nodes in discovering more microscopic metastases was not
the question and indeed was never in doubt: the harder you look the more you see.

The textual analysis reveals potentially important information. The authors of two studies state a prior
conviction concerning the value of MLND.[12,26] There are sources of potential bias in these trial
reports which are summarised in Table 2. In particular, in three of the five do not provide an intention
to treat analysis and significant numbers of patients were excluded post-randomisation. In the other
two reports it was not clear whether there was an intention to treat analysis and in Wu et al [15] there
was >10% imbalance between the two arms, which was not explained.

The clinical context has changed over time. Four out of five trials predate the routine use of PET CT
scanning in the pre-operative staging of patients with NSCLC. No authors mention the use of post-
operative adjuvant chemotherapy which is considered standard for those with Stage III disease. So
any conclusions drawn are less applicable to current practice.

The assessment of risk of bias (Table 2) shows that there are methodological uncertainties for all the
studies. Of particular concern is the lack of intention to treat analysis in three of them and uncertainty
about it in the other two. There are few randomised studies of the effectiveness of surgery in lung
cancer and the RCTs which we have found and analysed here show poor reliability. Four of these
RCTs were included in a previous meta-analysis reported in late 2014.[14] We have added a fifth
study and performed a detailed analysis of the text and the data. A further meta-analysis including four RCTs and eight non-randomised studies has been completed. The limitations we have indicated above have not been overcome.[29] The claimed survival benefit from mediastinal dissection is not supported by reliable evidence and ideally its overall value should be tested in a large pragmatic randomised trial involving contemporary diagnostic, surgical and oncological practice as has been proposed as a trans-Atlantic collaboration.[6] It would have to run by an independent clinical trials unit. Until and unless the results of such a trial are available, patients should be made aware of the risks and benefits of each of the approaches and participate in a shared decision making discussion with their physician/surgeon on the best option for their individual situation. The authors are willing to work towards setting up such a trial and between us we have a track record in being involved in and leading multicentre clinical trials of oncology and surgery.
Table 1: Trialists starting position and conclusions

<table>
<thead>
<tr>
<th>First author</th>
<th>Start</th>
<th>End</th>
<th>Starting position</th>
<th>Authors’ Interpretation of the results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Izbiki</td>
<td>1989</td>
<td>1991</td>
<td>“To what extent [MLND] contributes to the chance of cure remains controversial.”[23]</td>
<td>“… [MLND] is a safe operation that can be performed with acceptable morbidity and mortality rates.”[23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>“[MLND] did not improve survival … hazard ratio 0.78 95% CI 0.47-1.24”[11]</td>
</tr>
<tr>
<td>Sugi</td>
<td>1985</td>
<td>1998</td>
<td>“… pulmonary resection without mediastinal lymph node dissection has been considered a palliative operation.”[12]</td>
<td>“… peripheral non-small-cell carcinomas smaller than 2 cm in diameter do not require [MLND].”[12]</td>
</tr>
<tr>
<td>Darling</td>
<td>1999</td>
<td>2004</td>
<td>“Unfortunately, despite the fact that surgical staging of mediastinal lymph nodes is thought to be important, most surgeons do not perform a complete lymphadenectomy at the time of lung cancer resection.”[26]</td>
<td>“…no difference in local (P = .52), regional (P = .10), or distant (P = .76) recurrence between the 2 groups.”[10]</td>
</tr>
<tr>
<td>Zhang</td>
<td>2006</td>
<td>2007</td>
<td>“Compared [MLNS], [MLND] carries the potential advantage of accurate staging and survival benefit. But it may also be associated with increased surgical risks by prolonging operation time, increasing blood loss, and resulting in more complications.”[15]</td>
<td>“[MLND] and [MLNS] have similar surgical risks and mediastinal staging effect in patients with NSCLC.”[15]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>“[MLND] had significantly better five-year survival than [MLNS] (55.7% vs. 37.7%, P = 0.005).”[15]</td>
</tr>
</tbody>
</table>

[MLND]: mediastinal lymph node dissection

[MLNS]: mediastinal lymph node sampling
Table 2: Risk of Bias Assessment based on information presented in the publications. (ITTA: intention to treat analysis)

<table>
<thead>
<tr>
<th>STUDY</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding</th>
<th>Incomplete outcome reporting</th>
<th>Selective outcome reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Izbicki[23]</td>
<td>Clear</td>
<td>Unclear</td>
<td>Not possible</td>
<td>Yes: No ITTA</td>
<td>No</td>
</tr>
<tr>
<td>Sugi[30]</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Not possible</td>
<td>Unclear</td>
<td>No</td>
</tr>
<tr>
<td>Wu [13]</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Not possible</td>
<td>Yes: No ITTA</td>
<td>No</td>
</tr>
<tr>
<td>ACOSOG[26]</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Not possible</td>
<td>Yes: No ITTA</td>
<td>No</td>
</tr>
<tr>
<td>Zhang[15]</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Not possible</td>
<td>Unclear</td>
<td>No</td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1
Flow chart of searches

Figure 2 a to d
Forest plots of comparison in meta-analysis.

A. Early mortality odds ratio
B. Late mortality hazard ratio
C. Local recurrence odds ratio
D. Distant recurrence odds ratio

Figures 3
Perioperative complications with Odds Ratio
Reference List


11. Izbicki JR, Passlick B, Pantel K, Pichlmeyer U, Hosch SB, Karg O, Thetter O: Effectiveness of radical systematic mediastinal lymphadenectomy in patients with resectable...


