Screening for lung cancer: a systematic review and meta-analysis of controlled trials

R L Manser, L B Irving, G Byrnes, M J Abramson, C A Stone, D A Campbell

Background: Lung cancer is a substantial public health problem in western countries. Previous studies have examined different screening strategies for lung cancer but there have been no published systematic reviews.

Methods: A systematic review of controlled trials was conducted to determine whether screening for lung cancer using regular sputum examinations or chest radiography or computed tomography (CT) reduces lung cancer mortality. The primary outcome was lung cancer mortality; secondary outcomes were lung cancer survival and all cause mortality.

Results: One non-randomised controlled trial and six randomised controlled trials with a total of 245,610 subjects were included in the review. In all studies the control group received some type of screening. More frequent screening with chest radiography was associated with an 11% relative increase in mortality from lung cancer compared with less frequent screening (RR 1.11, 95% CI 1.00 to 1.23). A non-statistically significant trend to reduced mortality from lung cancer was observed when screening with chest radiography and sputum cytological examination was compared with chest radiography alone (RR 0.88, 95% CI 0.74 to 1.03). Several of the included studies had potential methodological weaknesses. Controlled studies of spiral CT scanning have not been reported.

Conclusions: The current evidence does not support screening for lung cancer with chest radiography or sputum cytological examination. Frequent chest radiography might be harmful. Further methodologically rigorous trials are required before any new screening methods are introduced into clinical practice.

Lung cancer is the commonest cause of cancer death in the western world.¹ It currently accounts for approximately 5% of all deaths in most developed countries and, as such, constitutes a major public health problem.¹ Current therapeutic interventions have had little impact on the epidemic proportions of the disease, and the case fatality rate remains at 85–90%.² The overwhelming majority of cases of lung cancer are attributable to cigarette smoking and thus primary prevention should continue to be a major focus of public health campaigns. However, such measures are likely to have only a limited impact on mortality in the short term because of a lag phase in the order of 20 years.

Most cases of lung cancer present at an advanced stage and therefore previous studies have investigated the role of screening for the detection of preclinical disease.¹ ³ Following a series of lung cancer screening trials conducted in the 1970s, it has generally been felt that early detection of lung cancer with chest radiographic or sputum cytology screening does not improve outcome, particularly disease specific mortality.³ However, recent narrative reviews have drawn conflicting conclusions.³ ⁴ Furthermore, newer screening methods have now been proposed such as spiral computed tomographic (CT) scanning.⁵ We therefore conducted a systematic review in order to assess the evidence for various screening methods to reduce lung cancer mortality and to evaluate the morbidity and harms associated with screening. We chose the primary outcome measure of disease specific mortality. This outcome is not influenced by the screening biases that may affect survival data—in particular, lead time and overdiagnosis bias. Lead time bias may occur if screening advances the time of diagnosis but does not actually alter the time of death, hence survival appears longer but mortality is unchanged. Overdiagnosis bias may occur if a screening programme detects cases of cancer that would not have led to death in that individual’s lifetime. This may occur if indolent cancers are detected, or if cancers are detected in individuals who would have died from co-morbid disease before the cancer became clinically apparent. If overdiagnosis bias occurs there may be an apparent improvement in stage distribution, resection rates, and survival without an improvement in disease specific mortality. Because the most appropriate outcome measure for screening trials is debatable, we have also considered the secondary outcomes of survival and all cause mortality.

METHODS

Identification of studies

Medline (1966–2000), Premedline (up to April 2001), Embase, and the Cochrane Controlled Trials Register were searched. The full search strategy is outlined elsewhere.⁶ We also contacted experts, examined bibliographies, and hand searched the journal Lung Cancer (1985–2000).

Selection of studies

Randomised or non-randomised controlled trials that examined the impact of screening for lung cancer in adult populations on lung cancer mortality were included. All screening interventions were considered including chest radiography, sputum cytological examination, and CT scanning. Studies that compared different screening modalities or different screening frequencies were included.

Two reviewers independently assessed the titles and abstracts from the electronic searches and relevant articles were selected for full text review. Studies were selected for inclusion in the review after both reviewers assessed the full text articles. Disagreements were resolved by consensus. When assessing the eligibility and quality of studies, the reviewers were aware of the authorship and source of publication of the studies.
Study quality

Study quality was evaluated independently by two reviewers and disagreements were resolved by consensus. For randomised controlled studies quality was assessed by noting the method of randomisation and whether allocation was concealed. For all studies we assessed whether there was blinding of the outcome assessment and whether there was an adequate description of withdrawals and dropouts. The criteria used to assess the adequacy of allocation concealment are described in the Cochrane Handbook.10 For the remaining quality criteria we used the descriptions outlined by Jadad et al.11

Data extraction

One reviewer extracted the data, and a second study member extracted the data for the main results. Authors of included studies were asked to confirm the data extracted. Some results were extracted from graphs.

Outcome measures

The primary outcome was lung cancer mortality. Secondary outcomes included lung cancer survival, all cause mortality, and morbidity associated with screening.

Statistical analysis

Outcomes from included trials were combined using the Review Manager (Version 4.1, Update Software, Oxford).12 For dichotomous outcomes, relative risks (RR) are reported with 95% confidence intervals (CI). Homogeneity of effect sizes between studies being pooled was tested using p<0.10 as the cut off level for significance. For those outcomes where there was significant statistical heterogeneity, relative risks were reported using the random effects model, but for other outcomes the fixed effects model was used. The fixed effects model considers only the within study variability in the calculation of the common effect whereas the random effects model takes into account both between study and within study variability in the calculation of the common effect. Random effects models will usually produce wider confidence intervals. Data were analysed on an intention to screen basis. The level of agreement between reviewers evaluating studies for inclusion and undertaking quality assessments was assessed using simple kappa and weighted kappa statistics.

Survival data were summarised using relative risks. We were unable to calculate hazards ratios for the survival analysis because the primary studies did not include sufficient information. None of the primary studies included life tables, Kaplan-Meier survival curves, or hazards ratios. Some of the studies did include graphs of cumulative event rates which could be extracted and used to approximate the hazards ratio. However, this method ignores censoring and effectively will approximate the ratio of cumulative mortality—that is, the risk ratio.

RESULTS

Literature search and inclusion in study

A total of 1869 citations were identified by the Medline search, 119 of which were selected for full text review (kappa=0.54; moderate agreement). Following the full text review, six studies (all with multiple citations) were selected for inclusion in the review (kappa=0.9; very good agreement). In four of the studies the investigators, who were masked to the screening status of subjects, assessed the cause of death.15 17 Withdrawals and dropouts were adequately described in four studies.1 4 14 15 In the Erfurt County study, however, losses to follow up were significantly greater in the control group (4.9% v 3.6%, p=0.0001).1 Follow up was poor in the Kaiser Permanente study14 and was not adequately reported in the Czech study.14 In the Johns Hopkins study 1.3% of participants were lost to follow up but no further details were provided.15 Extended follow up has recently been reported for two of the studies.11 19 In the Mayo Lung Project vital status was ascertained by searching the National Death Index,20 but the methods of follow up were not described in the Czech study.21

Compliance with screening

In the Mayo Lung Project compliance with scheduled screening averaged 75% in the intervention group and 73% of the control group received non-study chest radiographs during the final 2 years of the study.1 In the North London study24 63.2% of workers in the intervention group and 62.7% of workers in the control group attended for the final radiograph at the end of 3 years. In the Erfurt County Study25 compliance with scheduled screening was not described in detail. In the Czechoslovakian study26 92% of the intervention group and 95% of the control group attended screening that took place at the end of the first 3 years. In the Kaiser Permanente study14 the mean number of multiphasic health check ups (MHCs) per person during the study period was 6.8 (maximum 18, median 6) in the intervention group and 2.8 (maximum 17,
Lung cancer mortality

For studies which effectively compared more frequent chest radiographic screening with less frequent screening, the relative risk (RR) was 1.11 (95% CI 0.95 to 1.31) and there was no significant statistical heterogeneity between the results of different studies (p=0.67). The analysis was also conducted using data from the extended follow up reported for two of the studies (table 2).

Lung cancer survival

Survival was examined by comparing the proportion of patients diagnosed with lung cancer who were alive 5 years after diagnosis. For studies which compared more frequent chest radiographic screening with less frequent screening, survival was better in the intervention group with a combined RR of death from lung cancer of 0.91 (95% CI 0.84 to 0.99; p=0.02). There was significant statistical heterogeneity between the results of studies being pooled (p=0.023). For studies that compared annual chest radiographic screening plus 4 monthly sputum cytological examinations with annual chest radiographic screening, survival was better in the group receiving regular sputum cytological examinations (RR 0.83 (95% CI 0.75 to 0.92; p=0.0003)).

All cause mortality

A pooled analysis (table 4) for all cause mortality was conducted on those studies which effectively compared more frequent chest radiographic screening with less frequent screening (RR 1.01 (95% CI 0.94 to 1.08)). There was significant statistical heterogeneity between the results of the studies (p=0.08). Visual inspection of the graph for overlap of the

Table 1 Design of screening studies comparing different frequencies of chest radiographic screening

<table>
<thead>
<tr>
<th>Study and year commenced</th>
<th>Subjects</th>
<th>Intervention</th>
<th>Control</th>
<th>Screening duration</th>
<th>Total follow up*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erfurt County (1972)18</td>
<td>Men aged 40–65 years. Smokers and non-smokers</td>
<td>6 monthly chest radiographs</td>
<td>Chest radiographs every 1–2 years</td>
<td>6 years</td>
<td>6 years</td>
</tr>
<tr>
<td>North London study (1960)14</td>
<td>Men aged 40–65 years. Smokers and non-smokers from 119 work sites (mainly factories)</td>
<td>6 monthly chest radiographs</td>
<td>Chest radiograph at baseline and at the end of 3 years</td>
<td>3 years</td>
<td>3 years</td>
</tr>
<tr>
<td>Czech study (1976)16</td>
<td>Men aged 40–64 years. Current heavy smokers.</td>
<td>6 monthly chest radiography and sputum cytology for the first 3 years followed by annual chest radiograph for 3 years</td>
<td>Chest radiograph at baseline and chest radiograph and sputum cytology after 3 years followed by annual chest radiographs for 3 years</td>
<td>6 years</td>
<td>Initially 6 years, later extended to 15 years</td>
</tr>
<tr>
<td>Mayo Lung Project (1971)4</td>
<td>Men attending the Mayo Clinic aged more than 45 years. Heavy smokers</td>
<td>4 monthly chest radiography and sputum cytology</td>
<td>Advised at the start of the study to have an annual chest radiograph and sputum cytology test</td>
<td>6 years</td>
<td>Initially 11 years, later extended to 24 years.</td>
</tr>
<tr>
<td>Kaiser Permanente study (1964)14</td>
<td>Men and women aged 35–54 years. Smokers and non-smokers. Members of Kaiser Permanente medical care programme</td>
<td>Encouraged to undergo an annual multiphasic health check up including an annual chest radiograph</td>
<td>Subjects not urged to undergo screening but could do so as part of their usual care if requested</td>
<td>16 years</td>
<td>16 years</td>
</tr>
</tbody>
</table>

*Follow up period includes period of active screening and post screening follow up. The maximum follow up is described; for some studies this varied as subjects were enrolled at different stages.

Table 2 Relative risk of death from lung cancer: studies comparing frequent chest radiographic screening with less frequent screening

<table>
<thead>
<tr>
<th>Study</th>
<th>No randomised</th>
<th>No of lung cancer deaths</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intervention</td>
<td>Control</td>
<td>Intervention</td>
</tr>
<tr>
<td>North London14</td>
<td>29723</td>
<td>25311</td>
<td>82</td>
</tr>
<tr>
<td>Czech study16</td>
<td>3171</td>
<td>3174</td>
<td>247</td>
</tr>
<tr>
<td>Mayo Lung Project4</td>
<td>4618</td>
<td>4593</td>
<td>337</td>
</tr>
<tr>
<td>Kaiser Permanente14</td>
<td>5156</td>
<td>5557</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>42668</td>
<td>38635</td>
<td>710</td>
</tr>
</tbody>
</table>

*Results were identical with random effects and fixed effects models.
95% confidence intervals suggested that the results of the Czech study differed from the other studies. When the Czech study was excluded from the analysis the RR was 0.97 (95% CI 0.94 to 1.01) and there was no significant statistical heterogeneity (p=0.47). All cause mortality was reported in only one of the studies comparing annual chest radiography with annual chest radiography plus 4 monthly sputum cytological examinations (RR 1.03 (95% CI 0.9–1.17)).

**Harms and morbidity associated with screening**

In general, harms associated with screening were poorly reported. Postoperative deaths were not reported in three of the studies. In the remaining studies the number of postoperative deaths was small and did not differ significantly between intervention and control groups; for most of the studies they appear to have been included as lung cancer deaths. Morbidity associated with diagnostic work ups and surgical procedures was not well described—for example, no details were provided on the proportion of subjects with positive screening results who required further invasive tests. The details of the therapeutic interventions received by participants—such as extent of surgery—were poorly described.

**DISCUSSION**

This is the first systematic review of lung cancer screening studies reported in the literature. A previously reported meta-analysis was not undertaken in the context of a systematic review. The results of the present meta-analysis suggest that, overall, more frequent chest radiographic screening does not result in reduced lung cancer mortality compared with less frequent screening. In fact, when data from the prolonged periods of follow up recently reported for two of the primary studies are included in the analysis, more frequent chest radiographic screening is associated with an 11% relative increase in lung cancer mortality compared with less frequent screening. Screening biases, such as lead time bias and overdiagnosis bias, will influence survival as an outcome in screening studies. The finding in this review of a significant increase in survival from lung cancer in association with an increase in disease specific mortality emphasises the unreliability of survival as an outcome measure in screening trials.

These outcomes differ in that survival is usually measured from the point of diagnosis and the denominator is restricted to those who are diagnosed with cancer. In contrast, mortality is assessed by evaluating the number of deaths among all those assigned to a particular intervention group during a defined time period. As highlighted in the introduction, mortality (either disease specific or all cause) is not affected by screening biases. A more detailed description of these biases is outlined elsewhere.

Screening with 4 monthly sputum cytological examination in addition to annual chest radiography was not associated with a reduction in lung cancer mortality compared with annual chest radiographic screening alone. However, the 95% confidence intervals were relatively wide and included a range of potentially clinically significant values—for example, the true effect might lie between a 26% relative reduction in lung cancer mortality and a 3% relative increase in lung cancer mortality. None of the studies in the review addressed the question of whether there might be a benefit from less frequent chest radiographic screening (with or without sputum cytological examination) compared with no screening. However, a multi-screening study which is currently underway has been designed to assess whether screening for lung cancer with annual chest radiography can reduce disease specific mortality compared with no screening.

We identified potentially important methodological weaknesses in most of the included studies. Not all the methods used to assess the quality of studies included in this review have been validated for screening studies. The adequacy of allocation concealment has been shown to be an important source of bias in obstetric treatment studies, but has not been evaluated extensively in other fields. Recent reports suggest that it may be an important source of bias in screening studies. It should be noted that, although inadequate concealment will usually favour the intervention arm, bias can occur in either direction. Inadequate randomisation could therefore give rise to imbalances in prognostic variables between intervention and control groups at baseline and this could account for the finding of increased mortality from lung cancer in the group undergoing more frequent chest radiographic screening. It is noteworthy, however, that adjustment for baseline prognostic factors in one study did not alter

---

**Table 3** Relative risk of death from lung cancer: studies comparing annual chest radiography with annual chest radiography plus 4 monthly sputum cytological examination

<table>
<thead>
<tr>
<th>Study</th>
<th>No randomised</th>
<th>No of lung cancer deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intervention</td>
<td>Control</td>
</tr>
<tr>
<td>Memorial Sloan</td>
<td>4968</td>
<td>5072</td>
</tr>
<tr>
<td>Johns Hopkins</td>
<td>5226</td>
<td>5161</td>
</tr>
<tr>
<td>Total</td>
<td>10194</td>
<td>10233</td>
</tr>
</tbody>
</table>

*With the random effects model the pooled results were 0.88 (95% CI 0.73 to 1.06)

**Table 4** Relative risk of death (all causes): studies comparing more frequent chest radiographic screening with less frequent screening

<table>
<thead>
<tr>
<th>Study</th>
<th>No in each group</th>
<th>No of deaths</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intervention</td>
<td>Control</td>
<td>Intervention</td>
</tr>
<tr>
<td>Erfurt County</td>
<td>41532</td>
<td>102348</td>
<td>3143</td>
</tr>
<tr>
<td>Czech study</td>
<td>3171</td>
<td>3174</td>
<td>341</td>
</tr>
<tr>
<td>Mayo Lung Project</td>
<td>4618</td>
<td>4593</td>
<td>688</td>
</tr>
<tr>
<td>Kaiser Permanente</td>
<td>5156</td>
<td>5557</td>
<td>585</td>
</tr>
<tr>
<td>Total</td>
<td>54477</td>
<td>115672</td>
<td>4757</td>
</tr>
</tbody>
</table>

*With the fixed effects model the pooled results were 0.98 (95% CI 0.95 to 1.02).
the results substantially.31 Of course, if randomisation was inadequate, there is little assurance that potential confound-
ers which were not measured or not known were balanced at baseline. The results of this review should therefore be inter-
preted with caution. In particular, the finding of an increase in lung cancer mortality in the group receiving more frequent
chest radiographic screening may be the result of systematic
errors in the primary studies.

Bias may also have occurred as a result of the misclassifica-
tion of lung cancer deaths. Deaths due to other causes might
be more likely to be attributed to lung cancer in the group
undergoing more frequent screening.32 This has been termed
"sticking diagnosis bias". It is also possible that the treatment
of "pseudodisease" could have led to a relative increase in the
number of treatment related deaths.33 In addition to the above
limitations, it could be argued that studies conducted several
decades ago cannot be generalised to current practice. In par-
ticular, women were excluded from most of the studies and
the incidence of different histological subtypes of lung cancer
has changed since the 1970s.26

We certainly need to consider the possibility that screening
with frequent chest radiographs might be harmful. Based on
current understanding of the effects of medical radiation, it
seems unlikely that the increased incidence and mortality of
lung cancer is related to radiation exposure.26 Another possible
explanation is that, in the absence of a benefit from early
detection and treatment, early diagnosis adversely influences
outcome. When cancer is diagnosed in patients who would
otherwise have remained asymptomatic for some time, or
until death from another cause, the prognosis may be
influenced by the fact that they are now aware of the diagno-
sis. The diagnosis of cancer is frequently associated with
adjustment difficulties or depression. Higher levels of psycho-
logical stress in patients treated for cancer have been shown to
inhibit cellular immune responses,32 and coping and emo-
tional distress have been shown to be independent predictors
of survival in lung cancer.32

This review was limited to a small number of controlled
trials conducted over two decades ago. We adopted the
Cochrane methodology and therefore uncontrolled trials were
excluded. Although such studies represent a lower level in the
evidence hierarchy, they can be an important source of knowl-
dge and some might argue that their exclusion is a limitation
with frequent chest radiographs.32

In conclusion, there is currently insufficient evidence to
support screening for lung cancer with any screening modal-
ity. Our results suggest that there may be a role for sputum
cytology as an adjunct to other screening methods, but this
warrants further evaluation. Some experts have suggested
that current public policies that discourage routine chest
radiographic screening should be reconsidered,32 and others
have proposed non-comparative study designs for the evalua-
tion of cancer screening tools.31 However, we cannot ignore the
potential harms associated with screening asymptomatic
individuals in the community. Frequent chest radiographic
screening clearly does not significantly reduce lung cancer
mortality compared with less frequent screening and, given
the risk of false positive test results, the overall impact of such
screening might be detrimental. Emerging screening tech-
nologies need to be evaluated in well designed studies before
mass screening programmes are adopted.

ACKNOWLEDGEMENTS

The authors thank the Iberoamerican Cochrane Center for assistance
with database searches and Dr Consol Serra (coordinator of the
Cochrane Lung Cancer Group) for assistance with protocol develop-
ment and editing of the review. They acknowledge the help provided
by authors of primary studies who have responded to our correspond-
ence and provided additional information (Dr Robert Fontana, Dr
Myron Melamed) and are grateful to Carol Roberts who assisted with
the literature search and the retrieval of studies relevant to the review.

References

1. Lopez A. The lung cancer epidemic in developed countries. In: Lopez
AD, Caselli G, Valkonen T, eds. Adult mortality in developed countries:
from description to explanation. Oxford: Oxford University Press,
2. Ries LA et al. Influence of disease, histology and demographic factors
on lung cancer survival in the SEER population-based data. Semin
3. Wilde J. A 10 year follow-up of semi-annual screening for early
detection of lung cancer in the Erfurt County, GDR. Eur Respir J
cancer. A critique of the Mayo Lung project. Cancer 1991; 67(3
Suppl): 1153–64.
action project: overall design and findings from baseline screening.
of randomised clinical trials: is blinding necessary? Control Clin Trials
1996; 17:1–12.
13. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform
meta-analyses of the published literature for survival endpoints. Stat Med
1998; 17: 2815–34.
cancer. Results of the Memorial Sloan-Kettering study in New York.
screening for cancer of the lung: follow-up report of a randomised
controlled trial on a population of high-risk males in Czechoslovakia.
Int J Cancer 1990; 45: 26–33.
LUNG ALERT

Integrated PET-CT improves accuracy in staging of NSCLC compared with PET and CT alone

Positron emission tomography (PET) is increasingly used in the evaluation of non-small cell lung cancer. However, PET is imprecise in defining abnormal anatomy and computed tomography (CT) is superior. Scanners have been developed that incorporate PET and CT within the same machine and integrate the images produced. This study compares the diagnostic accuracy of this method with conventional techniques.

Conventional staging was performed on 50 patients using bronchoscopy and contrast enhanced CT scans; all had PET and integrated PET-CT scans. Forty patients underwent surgery; tumour stage was confirmed histologically in all patients and the nodal stage in 37. Two independent review bodies prospectively analysed the images. The results showed that, for tumour staging, integrated PET-CT scans were more accurate than CT alone (p<0.001), PET (p<0.001), and visual correlation of CT and PET scans (p=0.013; a value of 0.017 was deemed significant following Bonferroni's correction for multiple comparisons). For nodal staging, integrated PET-CT scans were better than PET scans (p=0.013); there was no significant difference compared with CT or visual correlation of scans.

This is a promising study that indicates that integrated PET-CT may be the most appropriate imaging for assessment of non-small cell lung cancer. Larger trials are required to confirm efficacy and assess cost effectiveness.

K Ryanna
Specialist Registrar, St Richard’s Hospital, Chichester, UK
kryanna@hotmail.com