CI

e-med
www.e-med.org.uk

ARTICLE

The pitfalls of lung cancer screening

M Goodwin* and **F V Gleeson**†

**John Radcliffe Hospital, Oxford, OX3 9DU, UK;* †*Churchill Hospital, Oxford, OX3 7LJ, UK*

Corresponding address: F V Gleeson, Churchill Hospital, Oxford, OX3 7LJ, UK
Tel.: +44-1865-225945; fax: +44-1865-225946; E-mail: `fergus.gleeson@radiology.ox.ac.uk`

Abstract

Lung cancer screening has received extensive attention for a number of years. As yet the goal of such a screening programme, a reduction in lung cancer mortality proven by a large randomised controlled trial, has not been achieved. Instead we are left with a number of unanswered questions and practical problems. In addition to the basic requirements for an effective screening programme, this review will identify the main pitfalls in lung cancer screening, with particular reference to multislice computed tomography. The specific difficulties relating to the identification of unimportant disease, the failure to identify important disease successfully, the consequences of investigating and treating identified disease and the financial costs will all be discussed.

Keywords: *Lung cancer screening.*

Introduction

Lung cancer is responsible for approximately one million deaths per year, with smokers making up 80% of these[1]. Surgical resection is currently the only hope of cure, but this can only be achieved in early stage disease. Unfortunately even in Stage I disease, the 5-year survival after surgery is about 70% and, more disappointingly, only 20% of lung cancers are currently diagnosed at Stage I[2]. The advanced nature of disease at diagnosis results in an overall 5-year survival of 10–15%.

With the scale of the disease and its grim prognosis, it is not surprising that a considerable body of work has gone into attempting to detect early stage disease using screening.

We are now able to screen for lung cancer using the chest radiograph (CXR), sputum cytology, fluorescent bronchoscopy, breath tests and multislice computed tomography (MSCT), but we are probably no closer to knowing whether we should screen than when the first CXR trials were initiated. The advent of single slice and then multislice computed tomography raised hopes of early detection and improved cure rates but unfortunately our ability to detect lung cancer earlier than ever using modern technology may not have produced the quantum leap in screening benefit first envisaged.

This article will briefly explain the important concepts in screening and use them to outline some of the inherent pitfalls, particularly related to the current use of MSCT in screening for lung cancer.

Basic principles of screening

Suggested criteria required for an acceptable screening test are given in Table 1. These will be briefly discussed with reference to lung cancer.

Table 1   Suggested criteria for a screening programme

(1)	High prevalence of disease
(2)	Low incidence of pseudodisease
(3)	The disease must be able to be detected in the pre-clinical phase
(4)	An effective test must be available to detect the disease in this pre-clinical phase
(5)	There must be an effective treatment for disease detected in the pre-clinical phase
(6)	The programme must be cost effective

The prevalence of lung cancer is certainly high and can be increased by selecting a high risk population to screen, for instance smokers over the age of 60 years. This effectively increases the sensitivity of the screening

test (because sensitivity depends on disease prevalence), and most of the reported studies on lung cancer screening (LCS) have applied this approach.

Pseudodisease is disease which does not affect the length or quality of a patient's life, and a large amount of pseudodisease within the screening population may render the programme ineffective[3]. Two types of pseudodisease occur. Type I pseudodisease is disease detected on screening which does not necessarily progress to symptomatic disease, for example, ductal carcinoma *in situ* of the breast, in which not all cases progress to invasive disease[4]. Type II pseudodisease is indolent disease occurring in patients who die from other causes; an example of this is prostate cancer, with more than 40% of men aged 60–70 years having histological evidence of disease[5], but only a small percentage of these having clinically evident disease. A significant amount of Type II pseudodisease in a screening population will result in overdiagnosis bias.

The concept of pseudodisease in lung cancer is controversial to some, but if present in sufficient magnitude would seriously diminish the likelihood of establishing a successful screening programme. For instance, is atypical adenomatous hyperplasia of the lung Type I pseudodisease, and are the increased numbers of peripheral ground glass adenocarcinomas detected on screening Type II pseudodisease?

The concept of a pre-clinical phase in lung cancer is important for the success of LCS. Fig. 1 depicts the natural history of a malignant disease amenable to screening. Screening for lung cancer can only work if it can be both detected and treated effectively before the critical point is reached, which in the majority of patients is the development of metastases. This means that screening must result in a shift in stage in detected lung cancer in order to reduce mortality.

The correlation between the size of the primary tumour, the presence of metastatic disease and prognosis seems self-evident initially, but the wide biological variability of malignant disease creates a more complex picture. One study of 510 Stage IA lung cancer patients demonstrated no significant correlation between tumour size and survival[6]. Another demonstrated tumour cells in the bone marrow of 55% of patients with T1 or T2 cancers[7]. These data suggest that simply detecting small lung cancers may not result in the expected reduction in mortality: the benefit of MSCT in lung cancer will be reduced if we detect apparently earlier stage disease judged by size, T stage, without a subsequent reduction in nodal and metastatic disease, N and M stage.

Biases other than overdiagnosis can occur in screening programmes, namely lead time bias and length time bias. Lead time bias occurs when screening results in an earlier diagnosis compared with usual practice, but the time of death remains unchanged: therefore, survival, the length of time the patient lives with the disease, increases in the screened population but mortality, death from the disease, is unchanged (Fig. 2).

Length-time bias occurs when screening detects slowly progressive disease which has a long pre-clinical phase and an assumed better prognosis compared with an aggressive disease which becomes symptomatic rapidly. An improved survival in the screened group then occurs due to the relatively higher proportion of indolent tumours detected in the screened group compared to the control group, whilst survival from aggressive disease remains unchanged.

It is also important to remember that although the test must be able to detect pre-clinical disease, radiological misses will occur—so called 'detection errors'—further reducing the effectiveness of the screening programme. In the Mayo Clinic LCS programme, four out of the 11 cancers detected on repeat screening had in retrospect been present on the previous scan, including one Stage IIIA tumour[8,9]. This inherent error rate bedevils breast screening radiology, leading in some instances to litigation, which has been in part responsible for difficulties in recruiting radiologists to breast screening programmes.

The potential pitfalls of LCS are now becoming recognized as the screening programmes world-wide report and analyse their findings and these are listed in Table 2. Each point will be discussed in turn.

Table 2   Potential pitfalls in LCS

(1)	The detection of unimportant incidental disease
(2)	The detection of important incidental disease
(3)	The detection of pseudodisease
(4)	Failure to detect important screened disease
(5)	The consequences of investigation and treatment of detected disease
(6)	Interval cancers
(7)	Opportunity costs

The detection of unimportant incidental disease

In the context of LCS, unimportant incidental disease predominantly relates to the presence of benign pulmonary nodules, usually granulomas. A number of LCS trials have now reported their initial results and all have demonstrated the difficulties caused by the large number of non-malignant nodules detected, which then require further investigation to confirm that they do not represent malignant disease[9–12,20,21].

The prevalence of pulmonary nodules in the screened population is high. In the Early Lung Cancer Action Project (ELCAP)[10,11] 23% of patients had non-calcified nodules at the initial prevalence CT. In the Mayo Clinic study 51% had non-calcified nodules at prevalence screening, rising to 69% of patients after 3 years of scanning. However, the prevalence of malignant nodules detected is much lower and has ranged from 1.3%[12] to 2.7% in the ELCAP study[10].
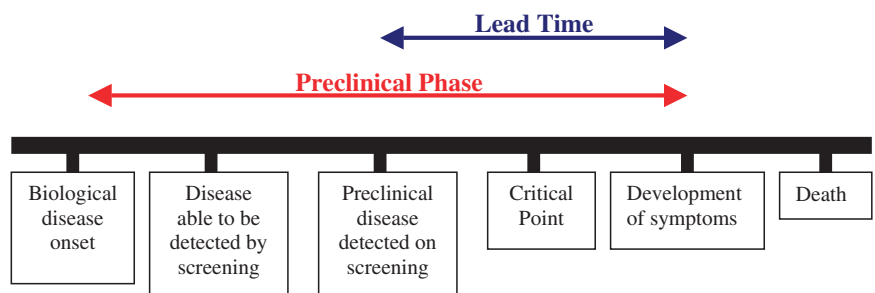
Figure 1   **A diagram demonstrating the different phases of a malignant disease.**
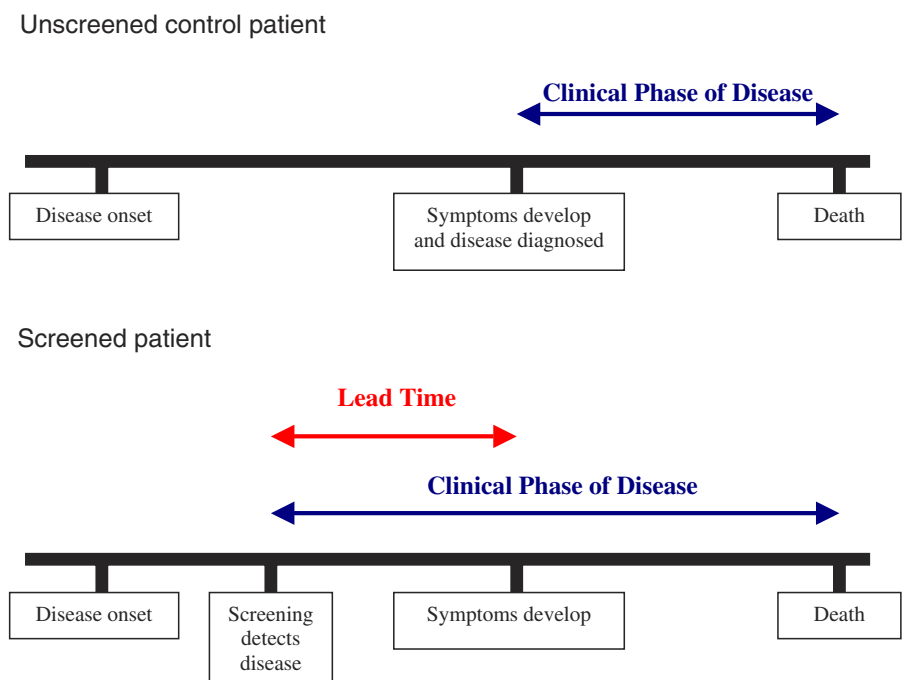

Figure 2   **This illustrates the concept of lead time bias.**

Further investigations are then required to differentiate between the benign nodules and the clinically relevant malignant nodules. The various protocols that have been developed are often complex, time-consuming and expensive. Coupled with such a large number of nodules to follow-up this represents another hindrance to an effective MSCT screening programme.

When nodules are detected on screening MSCT, the first task is to analyse the number of nodules and their morphology. More than six, or densely calcified nodules in a single patient are reported as granulomas and are considered to be of no importance. Less than six non-calcified nodules warrant further investigation by standard dose CT of the chest including volumetric thin section CT either through the whole chest or through the nodules. Nodule morphology is then analysed, and various imaging and interventional options are then used to determine their aetiology, depending on the research group's protocol. These options include follow-up scans, biopsy, assessment of the dynamic enhancement pattern, a positron emission tomography (PET) scan, or resection.

Most commonly the nodules are scanned at intervals to assess growth, which if present is considered to be highly suspicious for malignancy. These follow-up scans are performed at variable intervals (often at 3 and 6 months initially), with some investigators recommending a course of antibiotics prior to repeat scanning, to help decrease the number of infective or inflammatory nodules[11].

Clearly the detection and further investigation of large numbers of incidental benign nodules leads to a significant increase in workload and cost, together with patient anxiety and morbidity.

The detection of important incidental disease

Screening with MSCT is also likely to identify incidental disease outside the chest and this will increase workload

and cost. The Mayo Clinic group scanned the abdomen as well as the chest in their CT protocol, and findings included almost 700 additional abnormalities with 114 abdominal aortic aneurysms, four renal cell carcinomas, 63 indeterminate renal masses, 56 adrenal masses, 21 hepatic masses and 28 breast nodules. All of these required further work-up to assess their clinical significance.

The detection of pseudodisease

Most radiologists were unaware of the existence of the entity atypical adenomatous hyperplasia (AAH) prior to the advent of CT screening for lung cancer, let alone that it is found in 2–3% of patients at autopsy, and in 8–10% of patients undergoing resection for lung cancer[13]. It is typically a focal lesion often 5 mm or less in diameter and may be reported variably by pathologists. There is as yet not enough information available to be confident on the incidence of AAH detected in LCS programmes nor of its significance, but it may be the thoracic version of a Type I pseudodisease.

To most physicians and patients, lung cancer is regarded as an aggressive and for the most part incurable disease. The recent report analysing data from a 20 year follow-up of patients from the original Mayo Lung Project (MLP) raises concerns voiced almost 20 years ago; namely that screening may detect lung cancers that are not of importance[14]. The screened group, even after a 20 year follow-up, had a higher incidence of lung cancer but unchanged mortality compared with the control group. This would suggest that some screened detected cancers behave differently from cancers presenting symptomatically. It is known that adenocarcinomas make up a larger proportion of cancers detected at screening than is present in clinically presenting lung cancers, suggesting that some screen-detected adenocarcinoma may be a different biological entity. At one extreme, in a series of screening detected lung cancers 31% of the tumours were well-differentiated adenocarcinomas: all these were Stage I and their mean volume doubling time was 813 days[15]. Are these slowly growing tumours a form of Type II pseudodisease?

Will this problem be made more complex by utilising even more sensitive CT technology? T1 adenocarcinoma in the lung of less than 2 cm in size has been subdivided into six histological subgroups (A–F), which are associated with different prognoses: in one series the 5-year survival of types A and B was 100%[16]. Modern high-resolution CT of adenocarcinomas can help to differentiate between histological subtypes, particularly with reference to the amount of ground glass opacification present: in small peripheral lung adenocarcinomas detected by screening, 94% of type A lesions appeared as pure ground glass opacities, whereas type D cancers were homogenous nodules with no ground glass opacification[17]. In another study the presence of a higher percentage of ground glass opacification was confirmed as a useful prognostic marker, with these lesions having a significantly improved survival[18]. This is further evidence of the variable biological nature of lung cancers, fuelling the argument for the presence of significant pseudodisease.

Failure to detect important screened disease

The initial trials in LCS make sobering reading for radiologists. In the original MLP, even with triple reading of CXRs (with the sole purpose of detecting malignancy), up to 75% of peripheral and 90% of central lung cancers were visible in retrospect on review of previous films, i.e. 'missed'[19]. Recent reports suggest that a lesser number of lung cancers are 'missed' using CT. Nevertheless up to a third of CT screen-detected cancers are visible in retrospect: in the Mayo Clinic experience, four out of the 11 cancers detected on repeat screening had been present in retrospect on the previous scan, including one Stage IIIA tumour. Characteristic misses include cancers that are predominantly ground glass in appearance or associated with scars[20]. Help may be at hand via computer aided detection (CAD) software programmes, although these will have their own downsides, such as cost, reliability and increased time requirements.

The consequences of investigation and treatment of detected incidental disease

Although the initial report from ELCAP raised the possibility of being able to exclude all patients with benign disease from undergoing unnecessary biopsy or thoracotomy, other groups have not been so successful. The Mayo Clinic group had eight patients who underwent thoracotomy (21% of surgical procedures resulting from LCS) for benign disease[9]. In another study from Vancouver[21] three patients, 20% of those undergoing lung resections, had unnecessary thoracic surgery for benign disease.

The use of contrast enhancement as part of a protocol for lung nodule assessment, as performed by Pastorino *et al.*[22], may reduce the incidence of unnecessary thoracic surgery but is in itself not a perfect test. The large multicentre study assessing nodule enhancement reported by Swensen *et al.* also included false positive results, with a sensitivity of 98% and specificity of 58% (using a threshold of 15 HU as significant enhancement)[23]. To try to exclude these false positives Pastorino *et al.* raised the threshold for calling a positive test. PET scanning as reported in the same cancer screening trial may also be of value but is yet another test with reported failings particularly in the assessment of small nodules and indolent disease.

Assuming that a particular nodule has been labelled as suspicious as a result of the above investigations, percutaneous biopsy is usually the next step in order to obtain a histological diagnosis. Unfortunately, the majority of nodules that require biopsy will be 1 cm or less in size. To biopsy these lesions will be more technically difficult than other biopsies in most radiologists' current practice and the expected consequence of this will be a lower sensitivity, despite the outstanding results reported from ELCAP.

Even if the detected nodule is malignant there will be a morbidity and mortality from thoracotomy. From one large study of 12,439 patients who had lung resections, in-hospital mortality was 3.8% after wedge resection, 4.2% after lobectomy and 11.6% after pneumonectomy[24]. And one critique of LCS suggests that the consequent slow but accelerated decline in lung function secondary to pulmonary resection in the screened patients may be a further cause of death[25].

Interval cancers

It is depressing that even in an intensive screening programme such as ELCAP some lung cancer patients present symptomatically between the screening rounds (so-called interval cancers). Prior to the first interval scan following initial screening, two patients presented symptomatically, both with endobronchial abnormalities on CT[11]: one of these cancers was a limited small cell carcinoma, and the other had Stage IIB non-small cell lung cancer (NSCLC) resected successfully. These were two of nine carcinomas diagnosed by the end of the first interval screening round.

The Mayo Clinic experience also documents interval cancers. Within their screening period there were 10 NSCLCs and one small cell cancer detected by CT, but there were two interval cancers in the same period: one of these was Stage IV NSCLC and the other was a small cell cancer.

These data raise concerns as the proportion of interval cancers is high compared to the screen-detected cancers, and although expected for a lung cancer presenting symptomatically, they are at a comparatively more advanced stage. This appears to be further evidence of the biological variability in disease and it would seem unlikely that the outcome of such highly aggressive cancers will be altered, even by an intensive screening programme.

Opportunity costs

Most actions in medicine have a consequence and this is obviously the case in LCS. The limited resources available for health care suggest that until a LCS programme is proved to be successful in reducing disease-specific mortality any such strategy might be a net user of resources. Indeed it is possible that even if it successfully reduced lung cancer mortality it would continue to consume resources. After all, most symptomatic patients with lung cancer only survive for a limited period, whereas the screened population are likely to require screening for life, along with the other identified additional costs such as PET scans, biopsies, and the investigation of incidental disease. Probably, at least in the United Kingdom, the screening programme would consume resources by taking them from elsewhere within the health care environment. It may be that a greater patient benefit would occur if the funding of a screening programme were to be utilised in an alternative manner.

Conclusion

The pitfalls in LCS are comparable to those in screening programmes already in place in medicine. These include the identification of unimportant disease, the failure to identify important disease successfully, the consequence of investigating and treating disease identified, and the expenditure of money that may be better utilised elsewhere. All of these issues would be better assessed following a prospective randomised trial of multislice CT, when true efficacy and cost benefit could be assessed.

References

[1] Strauss GM. Screening for lung cancer. Surg Oncol Clin N Am 1999; 8: 747.
[2] Porter JC, Spiro SG. Detection of early lung cancer. Thorax 2000; 55(Suppl 1): S56.
[3] Henman CR, Gill HK, Eng J, Fajardo LL. Screening for preclinical disease: test and disease characteristics. Am J Roentgenol 2002; 179: 825.
[4] Yen MF, Tabar L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. Eur J Cancer 2003; 39: 1746.
[5] Montie JE, Wood DP Jr, Pontes E, Boyett JM, Levin HS. Adenocarcinoma of the prostate in cytoprostatectomy specimens removed for bladder cancer. Cancer 1989; 63: 381.
[6] Patz EF Jr, Rossi S, Harpole DH Jr, Herndon JE, Goodman PC. Correlation of tumor size and survival in patients with stage IA non-small cell lung cancer. Chest 2000; 117: 1568.
[7] Pantel K, Izbicki J, Passlick B *et al*. Frequency and prognostic significance of isolated tumour cells in bone marrow of patients with non-small-cell lung cancer without overt metastases. Lancet 1996; 347: 649.
[8] Swensen SJ. CT screening for lung cancer. Am J Roentgenol 2002; 179: 833.
[9] Swensen SJ, Jett JR, Hartman TE *et al*. Lung cancer screening with CT: Mayo Clinic experience. Radiology 2003; 226: 756.
[10] Henschke CI, McCauley DI, Yankelevitz DF *et al*. Early lung cancer action project: overall design and findings from baseline screening. Lancet 1999; 354: 99.
[11] Henschke CI, Naidich DP, Yankelevitz DF *et al*. Early lung cancer action project: initial findings on repeat screening. Cancer 2001; 92: 153.
[12] Diederich S, Wormanns D, Semik M *et al*. Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. Radiology 2002; 222: 773.

[13] Vazquez MF, Flieder DB. Small peripheral glandular lesions detected by screening CT for lung cancer. A diagnostic dilemma for the pathologist. Radiol Clin N Am 2000; 38: 579.

[14] Marcus PM, Bergstralh EJ, Fagerstrom RM *et al*. Lung cancer mortality in the Mayo Lung project: impact of extended follow-up. J Natl Cancer Inst 2000; 92: 1308.

[15] Hasegawa M, Sone S, Takashima S *et al*. Growth rate of small lung cancers detected on mass CT screening. Br J Radiol 2000; 73: 1252.

[16] Noguchi M, Morikawa A, Kawasaki M *et al*. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. Cancer 1995; 75: 2844.

[17] Yang ZG, Sone S, Takashima S *et al*. High-resolution CT analysis of small peripheral lung adenocarcinomas revealed on screening helical CT. Am J Roentgenol 2001; 176: 1399.

[18] Takashima S, Maruyama Y, Hasegawa M *et al*. Prognostic significance of high-resolution CT findings in small peripheral adenocarcinoma of the lung: a retrospective study on 64 patients. Lung Cancer 2002; 36: 289.

[19] Muhm JR, Miller WE, Fontana RS, Sanderson DR, Uhlenhopp MA. Lung cancer detected during a screening program using four-month chest radiographs. Radiology 1983; 148: 609.

[20] Li F, Sone S, Abe H, MacMahon H, Armato SG 3rd, Doi K. Lung cancers missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. Radiology 2002; 225: 673.

[21] McWilliams A, Mayo J, MacDonald S *et al*. Lung cancer screening: a different paradigm. Am J Respir Crit Care Med 2003; 168: 1167.

[22] Pastorino U, Bellomi M, Landoni C *et al*. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. Lancet 2003; 362: 593.

[23] Swensen SJ, Viggiano RW, Midthun DE *et al*. Lung nodule enhancement at CT: multicenter study. Radiology 2000; 214: 73.

[24] Romano PS, Mark DH. Patient and hospital characteristics related to in-hospital mortality after lung cancer resection. Chest 1992; 101: 1332.

[25] Reich JM. Improved survival and higher mortality: the conundrum of lung cancer screening. Chest 2002; 122: 329.