Clinical Issues in the Management of Early Lung Cancer

James L. Mulshine

Abstract

Lung cancer is commonly diagnosed after metastatic spread, when therapies are rarely curative, providing an impetus for continued research directed at exploring approaches for cost-effective early lung cancer detection. Recently published pilot studies across three continents support a benefit of spiral computed tomography (CT) in detecting earlier stage non–small cell lung cancer. Improved resolution of early lung cancer is a result of significant changes in CT imaging hardware and software. The status and implications of these developments are reviewed. Many aspects of the management of screening for early lung cancer could be informed by optimizing the downstream clinical management of potential lung cancers identified by CT screening. The first and most critical issue is whether or not this improved detection rate is clearly associated with a reduction in lung cancer–related mortality. However, other related issues such as cost-benefit evaluations are also considered. If smaller, truly localized primary cancer can be routinely detected, then options for less morbid interventions would also be desirable. The rapid improvement in resolution and cost of spiral CT has provided a powerful impetus to reconsider the possibilities for achieving safe, economical, and meaningful early lung cancer detection.

Lung cancer is the world’s leading cause of cancer death because it is routinely diagnosed after metastatic spread (1). Currently available therapies are rarely curative in the setting of disseminated disease (2). Progress with tobacco control programs in the developed world are more than offset by growing tobacco promotion and usage in the developing world. Furthermore, in nations such as the U.S. with mature tobacco control efforts, the benefits to the millions of former smokers in terms of improved cardiovascular outcomes is undercut by the long-term carcinogenic consequence of tobacco exposure. More lung cancers are now being found in former than in current smokers (3). Because improvements in treatment of advanced lung cancer have been so modest, early lung cancer detection has been the focus of considerable research interest (4).

Regarding the early experience with scanning, it is important to point out that chest X-ray is a low sensitivity detection tool for early lung cancer. The limitations of the chest X-ray to routinely detect localized cancer have been evident in a number of prospective randomized trials, as recently reviewed by the U.S. Preventive Services Task Force (5). Even when chest X-ray screening was done every 3 months, a significant percentage of cancers were detected outside of the screening interval due to the development of symptomatic cancers or were unresectable at first detection. In contrast, as recently summarized, a number of pilot studies across three continents suggest promising potential with using spiral computed tomography (CT) to screen for asymptomatic lung cancer in high-risk cohorts (2). These trials have used different imaging as well as different clinical management approaches but have achieved fairly consistent results. The reports from these studies suggest that a consistent majority of detected cases are surgically resectable with a low frequency of symptom-detected cases (6–9). Long-term survival or cure, in general, is seen only with early, asymptomatic lung cancer, so that the reported rates of detection of stage I lung cancer, outlined in Table 1, are favorable developments.

Based on these promising early reports with spiral CT-based early lung cancer detection, the question is whether spiral CT is the tool to change lung cancer mortality outcomes. The validation of CT-based lung cancer screening depends upon the demonstration of improvement in lung cancer–related mortality and major trials are currently under way to address this point. This question matters not only because of the profound public health consequences, but because the remarkable progress with high-resolution imaging and image processing is providing new tools that may redefine approaches to screening for a chronic disease in a public health setting (10).

Multidetector CT Scanners

From 1998 to the present, several generations of multidetector CT scanners have been introduced, going from single detection scanners to instruments with 64 rows of detectors by late 2004. Both the time required to complete a study and the thickness of the acquired images have improved so that an imaging study of the entire thorax with 0.625 mm slice thickness can now be acquired over a period of several seconds. This small slice thickness allows for markedly better resolution image reconstruction than a decade ago. The 3-second imaging acquisition time for the entire thorax also further reduces the
problem of image degradation related to respiratory motion. A practical implication of this remarkable refinement in imaging resolution is that certain functions have become more feasible. One example is the development of robust programs to permit reliable computer-aided detection (CAD)–based diagnosis of lung nodules at a defined size threshold. The Food and Drug Administration has favorably reviewed R2 Technology, Inc.’s submission of a CAD program, ImageChecker, for the detection of suspicious lung nodules. With the introduction of a new class of 64-row CT scanners, some measure of CAD performance will be necessary to enable radiologists to efficiently review and interpret the vast amount of data offered by these types of studies.

### Computer-Aided Diagnosis in Breast and Lung Cancer Screening

From a radiologist’s perspective, the large number of images acquired by these new machines in the screening setting presents a daunting challenge. In the current environment for breast cancer screening, reimbursement to radiologists for screening studies is limited, at the same time that adverse malpractice liability rulings impose sharp financial risks (11, 12). Rapid development of robust image processing and computer-aided diagnostic tools is an urgent priority for the lung cancer screening setting, because the subject numbers along with volume of image data per subject represent an overwhelming data load.

It is evident that many aspects of the management of early lung cancer screening could be informed by a more mature public health approach to early breast cancer screening (13, 14). The key aspect of this approach is developing economical, safe, and effective approaches to identification of presymptomatic disease in otherwise healthy but high-risk populations. Although aspects of this lung imaging dilemma mirror the situation with mammographic imaging of breast cancer, there are also important differences. Mammography is a planar imaging technique. Through breast compression, the three-dimensional anatomy of the breast is imaged and represented in only two dimensions. Thus, it is not surprising that CAD techniques have had only slight impact on mammography (15). However, high-resolution CT imaging can render a much more faithful three-dimensional representation of tissue structure (16). For these reasons, CAD analysis may enable much more significant contributions to lung cancer screening than has been the experience with breast CAD.

The ability to precisely and reliably segment particular structures within an imaging study is a primary challenge of current CAD development. A major goal with these new image processing tools is to reliably evaluate small nodules with a precision exceeding that of human vision, but the challenge will be to routinely distinguish them from proximal adjacent normal structures such as blood vessels or pleura (17, 18). As is true for mammography, lung cancer screening will entail evaluating huge numbers of imaging studies, with a true positive rate typically of about 1% or less. For refinement of efficient data reduction, work flow managing software system will be essential to leverage the productivity of thoracic radiologists so that they can economically and reliably handle the anticipated lung cancer screening volume. From a system engineering perspective, the dialogue about CAD, image processing, reimbursement, professional liability and radiology manpower is much more advanced in the field of breast cancer screening. Thus, the field of lung cancer screening could benefit from analysis of the experience in implementation of breast cancer screening services. However, the technological innovations with lung cancer screening over the last 10 years have occurred at a much more rapid pace than corresponding changes in mammogram technology. A significant aspect of this progress has been the emergence of CAD capabilities to assist the radiologist in handling the massive amount of imaging information that is acquired with a high-resolution CT screening study generated by the new multidetector scanners (18).

### Cost-Benefit Issues with Spiral CT Screening

The importance of determining the most efficient diagnostic workup approach to evaluate a suspicious lung nodule has been highlighted in a recent paper on cost-effectiveness (19). In that paper, investigators from Cornell analyzed the economics of their screening management. The group from Cornell has had the most extensive experience with lung cancer screening of all institutions (6, 7, 20, 21). One of the distinctive aspects of their approach to screening emphasized the use of noninvasive case validation as part of an intentional effort to decrease the cost and morbidity of screening management. Furthermore, if smaller, truly localized primary cancer can be

<table>
<thead>
<tr>
<th>Institution</th>
<th>Number of subjects</th>
<th>Number of CT-detected lung cancers</th>
<th>Mean/median primary cancer size (mm)</th>
<th>Frequency of detecting postsurgical stage I tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornell (prevalence)</td>
<td>1,000</td>
<td>27</td>
<td>−14/≤10</td>
<td>0.85</td>
</tr>
<tr>
<td>Cornell (incidence)</td>
<td>1,000</td>
<td>7</td>
<td>12/8</td>
<td>0.71</td>
</tr>
<tr>
<td>Mayo Clinic (prevalence)</td>
<td>1,520</td>
<td>29</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mayo Clinic (incidence)</td>
<td>1,000</td>
<td>21</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>National Cancer Center (prevalence)</td>
<td>1,611</td>
<td>14</td>
<td>20/20</td>
<td>0.79</td>
</tr>
<tr>
<td>National Cancer Center (incidence)</td>
<td>770</td>
<td>22</td>
<td>15/20</td>
<td>0.82</td>
</tr>
<tr>
<td>Milan (prevalence)</td>
<td>1,035</td>
<td>11</td>
<td>21/NA</td>
<td>0.55</td>
</tr>
<tr>
<td>Milan (incidence)</td>
<td>996</td>
<td>11</td>
<td>15/NA</td>
<td>1.00</td>
</tr>
</tbody>
</table>
routinely detected, then less morbid surgical as well as other intervention approaches can be considered, as outlined at a National Cancer Institute (NCI) State of the Science Meeting (http://www.webtie.org/SOTS/html/LungHome.htm). In addition, the development of adjunctive tools, such as effective chemoprevention, becomes more critical. It is clear that optimizing all aspects of the clinical management of screen-identified lung cancer is an ongoing challenge in this new field.

Other investigators have reported much greater cost projections with the process of lung cancer screening (22). In this report, which entailed a sophisticated model approach to determining costs, the cost for quality adjusted life-year saved was in excess of $100,000. This report generated considerable attention and involves important new modeling approaches, yet the conclusions of this paper must be evaluated carefully. The “estimates” used for critical variables such as operative mortality were considerably higher than the “observed” operative mortality input from the Cornell screening experience. In recent publications from the Cornell and Milan groups, the frequency of diagnostic thoracic surgical procedures is much lower than in other published reports (22). These clinical management measures have considerable impact with regard to the cost and morbidity of screening (8, 23). In a recent discussion of the process of cancer screening, it also was noted that how the screening was organized has major implications for the ultimate cost of the service (24).

The potential benefit of CAD analysis is not limited to more efficiently managing the massive data load associated with high-resolution imaging but may also improve the precision and robustness of the clinical management of early lung cancer in the screening setting. Given the potential costs, economic and medical, related to complications of lung cancer management, it is essential that robust software tools developed for CT image analysis be subjected to rigorous validation. Because these tools are emerging so rapidly, there is a need to better educate the medical community as well as the public on the strengths and limitations of computer-aided techniques. So there are many aspects of the management of screen-detected lung cancer that can be optimized to potentially reduce the cost as well as morbidity of this approach. Definitive cost-effectiveness analysis of lung cancer screening awaits clear evidence for the utility of this approach, but in the meantime, there are many important research questions to investigate which may improve the likelihood of favorable screening impact.

Clinicians can assist in the refinement of these tools by participating in trials that are accumulating the CT images/clinical outcomes databases. These databases can be used to enable imaging tool validation. One example of this type of research is the NCI-sponsored Lung Image Database Consortium to accelerate the maturation of the image-processing tools (25). This is a group of academic investigators committed to creating a large, well-characterized database of images and clinical outcome data for CAD algorithm and validation research for early lung cancer detection. Tracking the growth of suspicious pulmonary nodules through time in a longitudinally monitored cohort may be one of the few reliable ways to determine the natural history of disease entities (26, 27). In addition, data mining of these archival imaging resources will allow elucidation of the critical features to enhance the specificity of nodule characterization. With spiral CT imaging, a new classification of lung cancer is emerging based on the CT appearance of the cancer. Through time, validated software algorithms can potentially allow these imaging tools to be applied in general care settings so that the quality and economy of the clinical management CT-detected lung cancer screening can be delivered in a more homogeneously, high-quality fashion.

Another critical issue is whether lung cancer screening evolves from research centers to more typical care centers, the same favorable outcomes will be seen. We and others have commented that variable results in terms of the morbidity and mortality of the surgical intervention are areas of concern in achieving optimal benefit from lung cancer screening (5, 28).

### Does Overdiagnosis Occur with Spiral CT Screening?

The slower progress in implementing lung cancer screening programs compared with breast cancer screening may in part arise from the longstanding controversy about potential overdiagnosis bias in the previous NCI chest X-ray screening trials. Overdiagnosis bias refers to clinical outcome reports that are not adjusted for disease that would remain clinically covert until death from other causes. In the process of cancer screening, a less virulent form of the disease may be detected that the individual would have naturally died with, rather than of. As in the previous chest X-ray screening trials, preliminary findings of a screening effect may be positive without ultimately manifesting any significant reduction in lung cancer mortality. Although there are other potential biases, the need to address overdiagnosis bias is the primary factor driving the initiation of randomized screening clinical trials using spiral CT. Clinical trials orthodoxy suggests that the only way to control for overdiagnosis in the setting of a new screening tool such as CT scanning is to determine if there is a significant reduction in lung cancer–related mortality in a CT-screened population compared with outcomes in an appropriate control population.

To this end, the National Lung Screening Trial has been initiated in the U.S. This trial will accrue 50,000 current and former smokers at coordinating centers throughout the country in a randomized comparison of chest X-ray versus spiral CT to evaluate their relative reduction of lung cancer–related mortality (http://www.cancer.gov/nlst). In the Dutch national trial, called the NELSON trial, a total of 20,000 high-risk individuals will be randomized to either CT screening or standard care. The use of annual chest X-ray is not considered standard of care in the Netherlands. As part of the NELSON trial, all subjects will undergo a 16-detector, high-resolution CT scan including computerized image processing to determine nodule volume as well as related computer-assisted diagnosis tools. These large randomized trials are largely designed around the initial imaging test used to detect an early lung cancer.

Typically, overdiagnosis bias would be considered to include cases of CT-detected small cancers that evolve in a benign fashion, where such “pseudo cancers” lack the biological aggressiveness to mediate lethality. This was purported to be the situation with the Mayo component of the lung cancer X-ray screening trials (29). However, the importance of overdiagnosis bias in the chest X-ray trials is now being reconsidered because of the concern about the methodologic flaws in the Mayo component of the NCI-sponsored trial (5). As suggested by
Fontana, the statistical power of the Mayo chest X-ray screening trial was eroded by the high rate of contamination, given the frequent chest X-ray evaluations in the control arm, coupled with low compliance in the experimental imaging arm (5, 30). Secondly, a recent retrospective evaluation of tumor doubling time based on chest X-ray evaluations of the Mayo screen–detected cases on the experimental arm concluded that the growth rate was generally consistent with clinically aggressive cancers (31). Finally, as reviewed by Humphrey, conclusive evidence of the occurrence of “benign” cancer like that found for prostate cancer with autopsy series has not been convincingly assembled for lung cancer (5).

Some have suggested that detecting and operating on small cancers is not useful, because the outcome with smaller cancers is not better than with larger primary cancers (32). Patz et al. reviews the experience of a referral center over the last two decades with routinely detected stage IA lung cancer and concludes that primary tumor size does not significantly influence mortality outcomes. The mortality outcome in this entire reported cohort was better than expected, with patients rarely having subcentimeter-sized primary cancers that were detected with ongoing screening efforts (32). This report is at odds with the preponderance of published reports that the frequency of metastatic involvement decreases with decreasing primary size (7, 33–39). Recently, this issue was carefully analyzed using the data from the NCI Surveillance, Epidemiology, and End Results program, and favorable outcome was found to be directly related to the size of the detected primary cancer (20). However, the frequency of nodal metastasis even with subcentimeter primary cancers remains around 10% with regional variation related to primary location, so we have not yet reached the most favorable size for finding all lung cancers before dissemination.

Although overdiagnosis remains an important theoretical concern as CT resolution continues to find ever smaller primary lung cancers, at this time, no firm evidence exists which allows a reliable estimate of its influence in lung cancer screening trials. Recent data suggest that small cancers share the virulence of larger primary lung cancers (20, 39). A panel of international expert pathologists evaluating resected primary tumors from all the cases of screen-identified lung cancer obtained on the International Early Lung Cancer Action Program study concluded that the histopathology is consistent with the invasive characteristics of conventionally detected lung cancers (http://www.ielcap.org/professionals/in_conf_7_sum.htm). In addition, a collaboration of clinicians from Milan and Cambridge analyzed the expression profile of a prospective series of primary lung cancer obtained in the course of a pilot lung cancer screening trial compared with a matched series of symptom-detected cancers (8). In that analysis, they found that CT-detected tumors were not distinguishable from symptom-detected tumors by cDNA expression analysis or by corresponding immunohistochemical analysis. Therefore, there are three lines of evidence—clinical, pathologic, and molecular—to suggest that screen-detected lung cancers in general behave in an aggressive fashion that is typical for symptom-detected lung cancer.

The concept of overdiagnosis may also be construed to include the situation where a clinically aggressive lung cancer is detected by CT screening but the patient expires first of a comorbid condition such as coronary artery disease, the major competing risk from smoking. In the decades since the last major NCI-sponsored lung cancer screening trials, there have been considerable improvements in cardiovascular outcomes. In addition, a critical but unappreciated change has been the influence of the competing risks for death in the large and growing segment of the U.S. population that has stopped smoking. During the 1970s, when the NCI-sponsored screening trials were being done, there were many more current smokers than former smokers. With progress in tobacco control, there are now close to 45 million former smokers, nearly equaling the number of current smokers (40, 41). As a result of these changing trends in morbidity and mortality, the number of tobacco-related deaths from lung cancer now greatly exceeds the number from heart disease (41, 42). The dominance of lung cancer as a determinant of health outcomes in cohorts of former smokers makes interpreting the significance of overdiagnosis in the current CT screening trials more straightforward than was the situation with the chest X-ray trials (42). In the setting of rapid improvement in the resolution of lung cancer imaging capabilities, the theoretical possibility of overdiagnosis is an important concern. It is of interest to note that in its most recent review, the U.S. Preventive Services Task Force did not find the available evidence for overdiagnosis convincing (5).

Validating and Advancing the Benefits of Screening

The explosive improvement in the resolution and cost of spiral CT has provided a powerful impetus to reconsider the possibilities with regard to finally achieving meaningful early lung cancer detection (10, 43). The enthusiasm for improved lung cancer outcomes related to rapid progress in high-resolution CT imaging and computer-assisted image processing techniques is well justified. However, many cautionary reports outline important concerns about the safety and value of screening as well. The oncologic research community needs to understand the nature of the challenges inherent in population-based screening and to develop thoughtful solutions that can objectively address these challenges. Many of the potential research questions in managing screen-identified lung cancer occur in care settings involving clinical specialists other than oncologists. However, there are other research questions where the trial methodologies and research structures used in oncology could have a major positive impact in addressing issues in the management of early lung cancer.

To realize any of this promise, however, there are many formidable challenges that need to be addressed in a population-based setting that are beyond the scope of pilot studies done in centers of excellence. The first and most critical issue is whether or not the improved detection rate of stage I lung cancer reported in these pilot screening studies is clearly associated with a reduction in lung cancer–related mortality. Definitive evidence of a lung cancer–related mortality reduction with high-resolution CT screening for lung cancer is critical before many major agencies will recommend lung cancer screening. Implementing research urgently to optimize diagnostic workup, surgical interventions, and other aspects of clinical follow-up is essential. This field is moving quickly, and the time to develop an evidence base for defining best clinical practices for lung cancer screenings is now.
Dr. Thomas Lynch: I remember seeing the paper in CCR [2004; 10:6023–8]: what I can't recall is whether the screen-detected cancers for which they found no difference in gene expression profile were of the ground glass opacity type or were bigger, more solid tumors. Would you predict that you might find differences between the GGO type cancer and the more solid nodule type cancer?

Dr. Mulshine: I don't think they found many GGOs in their series. Yes, there would be differences, and the better the gene profiling is, the more differences one would see. They did not use Affymetrix platforms; they used the spotted arrays from Oxford Gene Technology. The GGOs are a relatively small fraction of all screen-detected cancers. So, even if you had 10% of them in a big cohort, they may not influence this kind of global assessment.

Dr. Bruce Johnson: In most of the array studies that have been published on lung cancer, they have to have 40% to 70% malignant cells within the tumor, so anything that wouldn't fit that profile wouldn't fit with what's typically published in the literature. Most of it is 70%. They do that by volume, so if you have enormous tumor cells and a bunch of small interstitial cells, the slough from the connective tissue, you may have a similar number of nuclei, but the assumption is that the mRNA would be greater from the bigger cells. That's the first point. The second is that from the two recent array studies [Proc Natl Acad Sci USA 2001;98:13790–5; Nat Med. 2002;8:816–24], there was an identifiable signature for tumors identified as having BAC features by one of three pathologists, so you should be able to pick those up with an expression array. The tools for comparing expression signatures across different platforms are very poor, so comparing with a non-Affymetrix platform is going to be tough.

Dr. Nick Thatcher: Computerized image analysis is clearly a big possibility, and we always talk about it, but I'm never clear how far along we are in routinely using it to help radiologists detect with better accuracy and efficiency. Where are we on that issue?

Dr. Mulshine: There has been one FDA submission so far, from R2 Technology in California. Their system has been approved and is commercially available but is definitely a first-generation offering. The challenge is that every spiral CT in the U.S. is replaced every 5 years, but to get the radiologists to buy the latest model, it has to have some feature that the radiologists want. Now the next-generation CTs are going to hit the market, and so CAD is an enormous push for the attractive to their client base. Unless they develop CAD, the next latest and greatest may not be so successful. The one caveat is that half or a little more than half of non–small cell lung cancer that goes to surgery performed and specialty board certification. So, this issue is hitting people—providers and patients—in their wallets, and so CAD is an enormous push for the manufacturers. Now, this is a good thing because the CAD tools will not only allow the radiologist to work more quickly, with more standardization, but these tools will potentially allow more reliable disease monitoring such that doing neoadjuvant window-of-opportunity drug trials will become more feasible because you would have a more quantitative and robust end point.

Dr. Claudia Henschke: In terms of the automated detection of nodules, I think we'll start seeing it in the next year and then more and more as better tools become available, but you do have to have the multiple slices and the high resolution. So, I think it will start coming into use, and I agree with Dr. Mulshine that it will increase the standardization. As for the growth rate tools, I think we really have to have some studies on how to best use them, what the pitfalls are, and then it should be integrated into the analysis.

Dr. Thacher: But that's not validated yet.

Dr. Henschke: We use it in our institution. We do it for other institutions, but it is misused. People send me slides where they say this showed that the cancer became smaller, and I don't know how they come up with that conclusion because you look at it and it's bigger. So there needs to be a lot of training and standardization.

Dr. Douglas Wood: We have been concentrating on the technology issues, but Dr. Mulshine also made some controversial comments about policy. I think all of us in this room would agree with the importance of specialization, and that comment has been made not just by Dr. Mulshine about the radiologists, but also by others regarding the pulmonary physicians or oncologists or thoracic surgeons who interpret what to do as the next step. Now, in this room we are preaching to the choir, and to an outsider our comments may seem self-serving, as if we're all saying that only high-level experts can manage the complexities of early stage lung cancer. So my question is, how do we affect policy so that we can have experts evaluating the CTs, making decisions on patients and treating them, because that is the way that we can best maximize the potential benefits of screening?

Dr. Lynch: I think you can market expertise. With computer-aided diagnostics and with people who become more sophisticated about the issues, I am optimistic that we will be able to get to the point where a 5-mm lesion can be managed effectively in the community.

Dr. Malcolm DeCamp: The one caveat is that half or a little more than half of non–small cell lung cancer that goes to surgery is not operated on by a thoracic surgeon.

Dr. Johnson: Is the outcome worse?

Dr. DeCamp: The outcome is worse looking in multiple administrative databases in terms of both volume of surgery performed and specialty board certification. So, this issue is hitting people—providers and patients—in their wallets, and that is the challenge if we try to disseminate this, not just to the academic community, but to the country.

References

Clinical Cancer Research

Clinical Issues in the Management of Early Lung Cancer

James L. Mulshine


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/13/4993s

Cited articles
This article cites 40 articles, 3 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/13/4993s.full.html#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
/content/11/13/4993s.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.