Chest X-ray Screening Improves Outcome in Lung Cancer*

A Reappraisal of Randomized Trials on Lung Cancer Screening

Gary M. Strauss, MD; Bay E. Gleason, PhD; and David J. Sugarbaker, MD, FCCP

It is believed that population-based screening for cancer should be advocated only when screening reduces disease-specific mortality. Four randomized controlled studies on lung cancer screening have been conducted in male cigarette smokers, and none has demonstrated reduced mortality. Accordingly, no organization that formulates screening policy advocates any specific early detection strategies for lung cancer. Yet, despite this public policy against screening, there is considerable evidence that chest x-ray screening is associated with earlier detection and improved survival. Two randomized trials, the Memorial Sloan-Kettering and Johns Hopkins Lung Projects, were specifically designed to evaluate the effectiveness of sputum cytologic study. Both evaluated the efficacy of the addition of sputum cytologic studies to annual chest radiographs, and both demonstrated that cytologic study did not favorably influence outcome. All individuals in experimental and control groups in both studies had annual chest radiographs. Because survival rates observed in both studies were about three times higher than predicted, based either on the National Cancer Institute's Surveillance Epidemiology and End Results database or based on the American Cancer Society's annual Cancer Statistics, raises the possibility that the periodic chest radiographs performed in all patients in both studies contributed to an improved outcome. In the Mayo Lung Project and in the Czechoslovak study on lung cancer screening, the experimental groups underwent a program of relatively intensive and regular rescreening with chest radiographs and sputum cytologic study, while the control groups underwent either less-frequent rescreening or no rescreening. In both studies, the screened groups achieved meaningful improvements in stage distribution, resectability, and survival. However, increases in cumulative incidence of lung cancer in the experimental group in both studies (which in the Mayo Lung Project reached statistical significance) prevented significant improvements in survival from translating into corresponding reductions in mortality. The possibility that screening may be associated with lung cancer "overdiagnosis" has been widely postulated to account for higher survival and incidence rates and equivalent mortality rates. However, analysis of autopsy information and of disease outcome in individuals with screen-detected early stage lung cancer who do not undergo surgical resection strongly supports the conclusion that screening does not lead to overdiagnosis of lung cancer. Similarly, lead-time and length bias do not adequately account for the differences in cumulative incidence observed in the Mayo and Czech studies. Because chest radiographs lead to meaningful improvements in stage distribution, resectability, and survival in lung cancer, and because neither overdiagnosis bias nor lead-time bias accounts for these improvements in outcome, a reconsideration of the role of chest radiographs in the early detection of lung cancer would be appropriate. A consensus conference would be a suitable forum to reexamine fully the existing data on lung cancer screening and to formulate specific guidelines for early detection strategies in individuals at high risk for lung cancer. (CHEST 1995; 107:270S-279S)

At the present time, it is widely accepted that screening for the early detection of lung cancer is not indicated.¹ This position is supported by the fact that no randomized controlled trial has succeeded in demonstrating that screening is associated with a reduction in lung cancer-specific mortality.

It has long been believed that disease-specific mortality, which correlates number of cancer deaths to the total number of individuals in the screened population, is the critical measure of cancer death rates. Unlike other outcome parameters, mortality is believed to be free of any confounding bias.² Mortality must be distinguished from fatality (or case fatality rate), the other measure of cancer death rates that is determined in the context of early detection trials. Fatality relates number of deaths to

*From the Division of Medical Oncology, Dana-Farber Cancer Institute, and the Division of Hematology-Oncology, Clinical Research Center, Endocrinology and Hypertension Unit, Department of Medicine, and Division of Thoracic Surgery, Brigham and Women's Hospital, and Harvard Medical School, Boston.

Supported by GCRC grant 5MO1RR02635.

Reprint requests: Dr. Strauss, Dana-Farber Cancer Institute, D1550A, 44 Binney St, Boston, MA 02115

270S

Multimodality Therapy of Chest Malignancies: Update '94
number of individuals with cancer. Because fatality is simply 100% minus percent surviving, fatality is not a unique end point. Rather, it is a different way of looking at survival.

All outcome parameters other than mortality can be confounded by bias. Accordingly, judgments regarding the effectiveness of screening based on such parameters as stage distribution, resectability, survival, and fatality can be misleading because of the potentially confounding effects of lead-time bias, length bias, selection bias, or overdiagnosis bias.

Because mortality is not thought to be subject to bias, it is considered an absolute standard that supersedes all other outcome parameters. If a statistically significant reduction in disease-specific mortality in a randomized study was achieved, there would be general agreement that the screening strategy used should become the standard of care.4

Our objectives are to demonstrate that existing studies on lung cancer screening show that periodic chest radiographs improve stage distribution, resectability, and survival, and that mortality is an applied measure of outcome that may not always accurately reflect cancer death rates in early detection trials. Based on substantial evidence showing meaningful clinical benefit, the role of annual chest radiographs deserves reconsideration by those organizations that make public policy on population-based screening.

**Randomized Trials on Lung Cancer Screening: Review of the Literature**

Four randomized controlled studies, which have collectively included 37,724 participants, have been conducted to evaluate lung cancer screening. Because these trials were all initiated in the 1970s, before the epidemic of lung cancer in women became so apparent, eligibility to participate in each of these trials was limited to male cigarette smokers. Although lung cancer subsequently surpassed breast cancer as the most common cause of cancer death in women,5 to our knowledge, no existing randomized trial has evaluated the efficacy of lung cancer screening in women. A problem with each of the four randomized trials is that none had a truly unscreened control group (Fig 1).

Three of the randomized trials were sponsored by the National Cancer Institute as part of the Cooperative Early Lung Cancer Detection Program.6 In the Memorial Sloan-Kettering Lung Project and the Johns Hopkins Lung Project, participants were randomized at study entry to a dual-screen group (in which subjects underwent annual chest radiograph and sputum cytologic study every 4 months) or to a single-screen group (in which annual chest radiographic screening was performed). In the Mayo Lung Project, all participants first underwent a prevalence screen, consisting of chest radiograph and sputum cytologic study. Those free of cancer were randomized to a study group (who underwent chest radiograph and sputum cytologic study every 4 months) or a control group (who received the recommendation for an annual chest radiograph and sputum cytologic study, but without efforts to enforce compliance).

In the Memorial Sloan-Kettering Lung Project (Table 1), 10,040 men were randomized.7 A total of 288 lung cancers were detected. More than 40% of the cases detected were in patients with stage I disease, and these patients with stage I disease achieved a 76% 5-year survival. Overall, 5-year survival for all patients was 35%. The total number of cancers, the number of late-stage cancers, the number of resectable cancers, and the number of cancer deaths were almost identical in the two groups. Similarly, in the Johns Hopkins Lung Project, there were no differences in outcome. Eight-year survival in both groups was 20%.8

These studies demonstrated no benefit by the addition of sputum cytologic study to annual chest radiograph. As randomized comparisons, these trials were designed to assess the benefit of sputum cytologic study, not of chest radiographs. However, each of the 20,427 men in these two trials underwent annual chest x-ray screening. Long-term survival in the Memorial and Hopkins studies was about three times better than that observed in the National Cancer Institutes' Surveillance Epidemiology and End Results (SEER) for men with conditions diagnosed during

<table>
<thead>
<tr>
<th>Table 1—Memorial Sloan-Kettering Lung Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual-Screen Group</td>
</tr>
<tr>
<td>Radiograph-Only Group</td>
</tr>
<tr>
<td>No. of patients</td>
</tr>
<tr>
<td>Lung cancers detected</td>
</tr>
<tr>
<td>Resectability, %</td>
</tr>
<tr>
<td>5-year survival, %</td>
</tr>
<tr>
<td>Lung cancer deaths</td>
</tr>
</tbody>
</table>

To our knowledge, no existing randomized trial has evaluated the efficacy of lung cancer screening in women. A problem with each of the four randomized trials is that none had a truly unscreened control group. Figure 1.
the same epoch. Thus, the two studies provide some evidence that the annual chest radiographs performed in all participants contributed to improved outcome, although such conclusions must be drawn with great caution. Selection bias is likely to confound the validity of any comparison involving different populations of patients, and other screening biases may also be operative.

Although the Memorial Sloan-Kettering and Johns Hopkins Lung Projects do not, as randomized studies, address the efficacy of chest x-ray screening, the other two randomized trials, the Mayo Lung Project and the Czechoslovak Study, provide some direct evidence of its efficacy. Although neither study had an untreated control group, both compared regular rescreening with periodic chest radiographs in an experimental group with infrequent, sporadic, or in some cases no rescreening in a control group.

The Mayo Lung Project\(^{10,11}\) is the most influential of the four existing randomized trials on lung cancer screening. All participants underwent a prevalence screen, which consisted of a chest radiograph and cytologic study before randomization. Patients found cancer free on the prevalence screen were then randomized to the experimental group, in which a chest radiograph and sputum cytologic sample were obtained every 4 months, or to the control group, which received the recommendation for an annual chest radiograph and cytologic study.

At every point when data were evaluated, the incidence of lung cancer was higher in the experimental than in the control population. At the study's end (Table 2), there were 206 cases of lung cancer detected in the experimental group compared with 160 cases in the control group. Resectability was higher in the experimental group, and lung cancer-specific 5-year survival was more than double, 33% vs 15%. However, despite this survival advantage, there were 122 lung cancer deaths in the experimental group and 115 in the control group. Accordingly, there was a statistically insignificant mortality advantage slightly favoring the control group.

The randomized trial from Czechoslovakia enrolled 6,364 men.\(^{12,13}\) All underwent a prevalence screen, consisting of a chest radiograph and sputum cytologic study. The experimental group then underwent chest radiograph and sputum cytologic study every 6 months for 3 years, while the control group underwent no screening until the end of the third year, when another chest radiograph and cytologic study were performed. Subsequently, both groups underwent chest radiography at the end of the fourth, fifth, and sixth years.

During the initial 3-year period (and excluding the scheduled screening at the end of the third year), 36 lung cancers were diagnosed in the experimental group and 19 in the control group (Table 3). Resectability was 25% in the experimental group and 15% in the control group. Survival was superior in the experimental group, with no survivors beyond 3 years among control subjects and 25% surviving 5 years in the experimental group, a difference that was highly statistically significant (p=0.0001). However, mortality was higher in the experimental group; there were 28 lung cancer deaths vs 18 in the control group.

During the course of the entire 6 years of the trial, a total of 108 cancers were detected in the experimental group and 82 in the control group. Overall, there were 85 lung cancer deaths in the experimental group and 67 in the control group. Accordingly, relative mortality was 27% higher in the experimental group, although this difference was not statistically significant (p=0.16).

Both the Mayo Lung Project and the Czech study (during the 3-year experimental period) showed advantages for the screened group with regard to stage distribution, resectability, and survival. However, there were statistically insignificant mortality disadvantages for the screened groups in both studies.

Furthermore, in both studies, the incidence of lung cancer was higher in the experimental group than in the control group. In the Mayo Lung Project, this difference in cumulative incidence was statistically significant (p=0.019), and in the Czech Study it approached significance (p=0.065). The question is why was this disparity observed? It is critical to understand the nature of these incidence discrepancies, because of the relationship between mortality and incidence. Cumulative mortality is proportional to two variables: cumulative incidence and fatality.\(^2\)

The numerator in a mortality calculation is the number of cancer deaths. The only candidates for lung cancer mortality are those persons with lung cancer, who constitute the incidence group. The
number of deaths then depends on the percentage of these patients who die of cancer, which is the definition of fatality.

**Potential Explanations for Differences in Cumulative Incidence**

The four screening biases represent potential explanations for these incidence discrepancies. Selection bias influences participation in screening trials, but randomization eliminates selection bias as a factor in whether individuals enter an experimental or control group.

Lead-time bias predicts an initial excess of cases, reflecting earlier detection through screening. This is due to increased detection of relatively indolent cancers, which, because of length bias, are most likely to be detected by screening. However, if all cancers are clinically significant lesions destined eventually to become life threatening if not effectively treated, then after screening is discontinued, the number of control cases should approach that of the experimental group.

In both the Mayo Lung Project and Czech study, however, the incidence discrepancy actually increased during a postexperimental period (Table 4). In the Mayo study, intensive screening in the experimental group continued for 6 years but was followed by a period of observation that, on average, lasted 3 years. At the end of the 6 years, there were 56 more cases in the screened group. During the first 18 months, this discrepancy narrowed to 31, but by the end of observation, it had increased again to 46.

The Czech study’s design was uniquely suited to eliminate lead-time bias. At the end of the experimental period, the incidence difference was 17, which narrowed to 9 when including those cases of lung cancer detected at the planned rescreening at the end of the third year. During the subsequent 3-year period when both groups underwent annual screening chest radiographs, the incidence discrepancy almost tripled to 26. Such incidence discrepancies, which increased during the postexperimental period when both experimental and control groups were treated, in effect, as a single screened population, are not explainable on the basis of lead-time or length bias. In the Mayo Lung Project, it is conceivable that lead time was so prolonged that it exceeded the screening and observation period, which averaged 9 years in both groups. However, if the lead time is so prolonged, it merges conceptually with over-diagnosis.

Over-diagnosis bias has been widely accepted as the most likely explanation for the observed incidence differences. For example, Eddy, in his review of lung cancer screening, concluded that the results of the Mayo Lung Project are “consistent with the hypothesis that many of the lesions detected by screening and labeled as cancers were not clinically important in the sense that they would never have become clinically evident during the time of the trial and follow-up.”

From a statistical and epidemiologic point of view, over-diagnosis in a screened group can plausibly account for presumed “missing cases” in a control group. However, it is highly questionable whether this hypothesis is biologically plausible. A malignancy as clinically virulent as lung cancer would not be a very likely candidate for over-diagnosis through screening. Independent of hypothetical considerations, however, evidence for over-diagnosis can be sought in two ways. The first is to determine if a substantial population with latent lung cancer detected incidentally at autopsy exists, similar to what is well known to occur with prostate cancer. It has been demonstrated many times that almost one third of men older than 50 years and two thirds older than 60 years who died of unrelated causes have evidence of latent prostate cancer at autopsy.

The only autopsy series that provides some insights into the phenomenon of latent lung cancer is from Yale. Among 3,286 autopsies, 159 cases of lung cancer were detected, which included 26 “surprise” cases completely unsuspected before death. This 0.8% incidence of surprise cases in the Yale autopsy series corresponds well to the presumed 1.0% incidence of “missing” cases in the Mayo Lung Project control group. Might these 26 surprise autopsy cases of lung cancer represent the same population as the missing cases in the Mayo control group? It should be kept in mind that all of the presumed Mayo Lung Project missing cases were cigarette smokers, with asymptomatic stage I lung cancer detected on routine screening chest radiograph and, of course, all were men.

The clinical profile of the surprise cases in the Yale autopsy series, however, was quite different. Among the surprise cases, stage distribution was skewed toward advanced-stage disease: 58% had regional or distant metastases. The radiograph was either normal or not suggestive of lung cancer in 91%. Women and nonsmokers were relatively more strongly represented among the surprise cases. Furthermore, 37% of autopsy cases were too ill for medical evaluation,
compared with only 2% of patients with disease diagnosed before death, suggesting that many patients with lung cancer diagnosed at autopsy appear to have died of lung cancer.

The second line of evidence relates to the hypothesis that if screening results in overdiagnosis of lung cancer, then many chest radiograph-detected lesions should remain clinically silent for prolonged periods, even if untreated. Consequently, a substantial number of patients not undergoing resection should be long-term survivors.

Again, prostate cancer provides the model. Chodak et al. recently reported a pooled analysis from six studies of 828 patients who were managed initially by observation without active therapy. Results demonstrated that 10-year prostate cancer-specific survival was 87% for men with grade 1 or 2 tumors.

Lung cancer and prostate cancer, however, are at opposite ends of the spectrum regarding malignant behavior. Flehinger et al. reported on the importance of resection in all stage I non-small cell lung cancer identified in the Mayo, Memorial, and Hopkins studies. Among 331 patients with stage I disease, 45 did not undergo surgery, either because of refusal or medical contraindication. Five-year survival for the 286 patients who were treated by resection was 70%. In contrast, lung cancer-specific 5-year survival for the 45 patients who did not undergo surgery was 10%.

These data are not consistent with the thesis that screening results in overdiagnosis of lung cancer. As Flehinger et al. pointed out, “Patients with lung cancer detected early who were treated by resection of their tumor had a high probability of survival,” while patients “with lung cancer detected early who were untreated died of lung cancer.”

Existing data support the conclusion that none of the four screening biases credibly accounts for the hypothetical missing cases in the Mayo and Czech studies. This is an important conclusion, because these are the factors that confound the significance of the survival advantages observed in these studies. If the survival advantages observed in the Mayo and Czech studies were not confounded by overdiagnosis, lead-time bias, or length bias, it would support the conclusion that screening was leading to meaningful improvements in outcome.

However, it also leads one to question what does account for the incidence discrepancies. Does some statistically or biologically meaningful explanation exist for the increased incidence in the screened groups that is not related to a screening bias?

One possibility is that radiation from screening chest radiographs was responsible for the observed increase in cases of lung cancer. However, cancer induction from the low doses of radiation is simply not plausible. Radiation exposure to the lung from a single posteroanterior and lateral chest radiograph is 0.02 rad. It has been estimated that, at most, the upper limit of lifetime risk from radiation exposures of 1 to 10 rad is one cancer death per 10,000 persons receiving I rad exposure, following a latency of 10 to 40 years.

Another possibility is chance alone. Assuming that the process of randomization resulted in two groups with equivalent lung cancer risk, chance accounted for incidence differences. That, however, is what statistical significance is all about. The possibility that chance alone produced differences in incidence in both the Mayo and Czech studies is the product of the p values of the incidence differences in the individual studies, or approximately one chance in 800 (p=0.019×0.0012). Accordingly, chance alone would represent an extremely unlikely explanation.

**IMBALANCE OF COEXISTING RISK FACTORS FOR LUNG CANCER**

We believe that existing data support the conclusion that neither the four conventional screening biases nor radiation exposure from screening chest radiographs nor chance can credibly explain the observed differences in cumulative incidence of lung cancer in the Mayo Lung Project and the Czechoslovak study. We believe it is necessary to look for an alternative explanation for the observed differences in case detection. We also understand that since none of the studies were designed to determine the nature of any unexpected incidence discrepancies, the answer to this question is speculative.

Nevertheless, we believe that the most plausible hypothesis for differences in cumulative incidence relates to an imbalance in covariates that determine lung cancer risk. This suggestion might be greeted with skepticism. Because cigarette smoking is the major factor leading to development of lung cancer in the vast majority of cases, how could other factors play a significant role?

In these studies, the role of smoking as the predominant etiologic agent in lung cancer is not relevant. All participants in both groups in the Mayo Lung Project and the Czechoslovak study were cigarette smokers, so there was no difference between groups with respect to this most critical of variables. However, there is a great deal of evidence that certain cigarette smokers are much more susceptible than others to development of lung cancer, which is dependent on coexisting risk factors (Table 5).

For example, the nature of the smoking history is important, including such variables as smoking duration and age at onset. Many occupational and environmental carcinogens affect lung cancer risk. A history of asbestos exposure increases the risk of lung
cancer fivefold among smokers,\textsuperscript{28} while a history of radon exposure increases it about threefold.\textsuperscript{29} The presence of chronic obstructive lung disease increases lung cancer risk 4.4-fold, based on a recent prospective study.\textsuperscript{30}

Regarding family history, there is a consistent pattern of increased lung cancer prevalence among close relatives of patients with lung cancer (from twofold to fivefold), even when adjusting for smoking habits.\textsuperscript{31,32}

Dietary factors appear to influence lung cancer risk. Many studies have demonstrated that certain antioxidant vitamins decrease lung cancer risk,\textsuperscript{33,34} although a recent randomized study suggested an increased risk of lung cancer associated with beta carotene.\textsuperscript{35} Finally, molecular genetic factors might be most influential in stratifying risk. Molecular epidemiology utilizes molecular genetic and biochemical techniques to evaluate risk cancer in individual patients.\textsuperscript{36} Within the next decade, it may be possible to develop an individual risk profile that integrates exposure to environmental carcinogens with such constitutional factors as altered proto-oncogenes and tumor suppressor genes.\textsuperscript{32}

It is absolutely and unequivocally true that cigarette smoking is the major determinant of risk for lung cancer. However, marked variation in individual susceptibility to development of lung cancer exists among cigarette smokers, depending on coexisting risk factors. If a large number of covariates are important in defining disease risk, an imbalance in the allocation of one or more of these variables could be responsible for variation in incidence, despite randomization.

If randomization did not result in groups with equivalent lung cancer risk, there is no reason to expect incidence to be equivalent. Because it is possible to exclude virtually all other hypotheses, an imbalance of risk-defining covariates would appear to be the most plausible explanation for inequality of incidence between experimental and control groups in the Mayo and Czech studies.

**Mayo Lung Project Reexamined**

In the Mayo study,\textsuperscript{10} the cumulative incidence of lung cancer in the experimental group was 4.46% vs 3.48% in the control group. The difference in absolute incidence was 0.98%. This led to a 29% relative excess in case numbers in the experimental population, indicating that the size of the group at risk for lung cancer mortality was 29% larger, a statistically significant difference (p=0.019).

If incidence rates between experimental and control groups in randomized trials differ, not because of a screening bias but because the groups differ with regard to their risk of developing lung cancer, then relative mortality rates would have no meaning. However, if overdiagnosis were not increasing case detection, all diagnosed cases would be clinically significant. In such circumstances, fatality rates more accurately reflect cancer death rates than mortality, since fatality is a measure of cancer death rates in the incidence group.\textsuperscript{2}

In the Mayo Lung Project (Table 6), there were 122 lung cancer deaths among 206 lung cancer cases in the experimental group, for a fatality rate of 59% vs 115 lung cancer deaths among 160 lung cancer cases in the control group, for a fatality rate of 72%. This difference in fatality is statistically significant (p=0.016).

**Changing Epidemiology of Lung Cancer**

In 1980, the American Cancer Society modified its previous recommendation in favor of annual screening chest radiographs in cigarette smokers to a recommendation supporting no screening whatsoever.\textsuperscript{37} Table 7 lists the changes that have occurred in lung cancer epidemiology from 1980 through 1994, based on the American Cancer Society's annual cancer statistics.\textsuperscript{38-52} During this period, annual incidence increased 47%. The percentage of women who contribute to the total pool of patients with lung cancer increased from 27 to 42%; furthermore, the number of women annually who are diagnosed as having lung cancer and die of lung cancer more than doubled.

However, this increasing incidence of lung cancer in our society has occurred despite a decreasing percentage of Americans who continue to smoke. This trend toward a decrease in smoking prevalence has taken place over the last several decades. In 1965, 42% of the population older than 18 years were ciga-

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Example</th>
<th>Increased Risk in Smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of smoking history</td>
<td>Age of onset, nicotine content, unfiltered cigarettes, inhalation</td>
<td>Variable</td>
</tr>
<tr>
<td>Environmental</td>
<td>Radon</td>
<td>3</td>
</tr>
<tr>
<td>Occupational</td>
<td>Asbestos</td>
<td>5</td>
</tr>
<tr>
<td>Benign lung disease</td>
<td>Emphysema and bronchitis</td>
<td>4.5</td>
</tr>
<tr>
<td>Family history</td>
<td>First- or second-degree relative with lung cancer</td>
<td>2.5</td>
</tr>
<tr>
<td>Dietary factors</td>
<td>Consumption of fruits and vegetables</td>
<td>Variable</td>
</tr>
<tr>
<td>Molecular genetic factors</td>
<td>Mutation or overexpression of oncogenes or tumor suppressor genes</td>
<td>Uncertain</td>
</tr>
</tbody>
</table>
are cigarette smokers. The percentage of smokers had fallen to 37% by 1974, 32% by 1983, and 29% by 1987. In 1991, the percentage of smokers had fallen to 26%.

Because lung cancer incidence is increasing at the same time that smoking prevalence is decreasing, lung cancer is becoming more and more a disease of former smokers. In 1991, it was estimated that 89.8 million US adults were at some point regular cigarette smokers. This includes a total of 46.3 million current and 43.5 million former smokers. The increasing incidence of lung cancer, coupled with the evidence that smoking cessation has not yet been an effective means of primary prevention, adds weight to the argument that the increased use of early detection methods will be necessary to reduce death rates from lung cancer.

**DISCUSSION**

Existing evidence from all four of the randomized controlled studies on lung cancer screening supports the conclusion that periodic chest radiographs lead to increased detection of early-stage disease, increased rates of resectability, and improvements in lung cancer-specific survival. A mortality reduction, however, was not demonstrated in any of these four studies.

In the Memorial Sloan-Kettering and Johns Hopkins Lung Project, the failure to demonstrate a mortality difference was likely due to the fact that both control and experimental groups in both studies underwent annual chest x-ray screening. Sputum cytologic study did not contribute to a further improvement in outcome.

In the Mayo Lung Project and the Czechoslovak Study, the two randomized studies that best address the efficacy of chest x-ray screening, the failure to demonstrate a mortality reduction was likely due to the higher cumulative incidence of lung cancer in the screened populations. While one might expect such a higher incidence because of the effects of lead-time, overdiagnosis, or length bias, the evidence indicates that these biases do not account for the higher cumulative incidence of lung cancer in the screened groups in either of these two studies.

The observed improvements in survival (as well as stage distribution and resectability), if not explainable by lead-time, overdiagnosis, or length bias, would lead to the conclusion that there is a striking benefit associated with screening. However, mortality comparisons would support no benefit. This paradox raises the question whether mortality comparison accurately reflected outcome.

Mortality would not be an accurate measure of cancer death rates when groups are not comparable with regard to their risk of the disease under investigation. This is true because mortality is proportional to the product of incidence times fatality. While this is subject of speculation, we believe that the most likely reason for the observed differences in cumulative incidence is that the groups were not balanced with regard to important variables that determine lung cancer risk.

Under such very limited circumstances, cancer

### Table 6—Fatality Rate Comparisons in Mayo Lung Project

<table>
<thead>
<tr>
<th>No. of Lung Cancer Deaths</th>
<th>No. of Lung Cancers</th>
<th>Fatality Rate, %</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>122</td>
<td>206</td>
<td>59</td>
</tr>
<tr>
<td>Control group</td>
<td>115</td>
<td>160</td>
<td>72</td>
</tr>
</tbody>
</table>

### Table 7—American Cancer Society Annual Estimates of Incidence and Deaths From Lung Cancer in the United States, 1990-1994

<table>
<thead>
<tr>
<th>Year</th>
<th>Incidence</th>
<th></th>
<th>Deaths</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Women</td>
<td>% Women</td>
<td>Total</td>
</tr>
<tr>
<td>1980</td>
<td>117,000</td>
<td>32,000</td>
<td>27.3</td>
<td>101,300</td>
</tr>
<tr>
<td>1981</td>
<td>122,000</td>
<td>34,000</td>
<td>27.9</td>
<td>105,000</td>
</tr>
<tr>
<td>1982</td>
<td>129,000</td>
<td>38,000</td>
<td>29.5</td>
<td>111,000</td>
</tr>
<tr>
<td>1983</td>
<td>135,000</td>
<td>41,000</td>
<td>30.4</td>
<td>117,000</td>
</tr>
<tr>
<td>1984</td>
<td>139,000</td>
<td>43,000</td>
<td>30.9</td>
<td>121,000</td>
</tr>
<tr>
<td>1985</td>
<td>144,000</td>
<td>46,000</td>
<td>31.9</td>
<td>125,600</td>
</tr>
<tr>
<td>1986</td>
<td>149,000</td>
<td>49,000</td>
<td>32.9</td>
<td>130,100</td>
</tr>
<tr>
<td>1987</td>
<td>150,000</td>
<td>51,000</td>
<td>34.0</td>
<td>136,000</td>
</tr>
<tr>
<td>1988</td>
<td>152,000</td>
<td>52,000</td>
<td>34.2</td>
<td>139,000</td>
</tr>
<tr>
<td>1989</td>
<td>155,000</td>
<td>54,000</td>
<td>34.8</td>
<td>142,000</td>
</tr>
<tr>
<td>1990</td>
<td>157,000</td>
<td>55,000</td>
<td>35.0</td>
<td>142,000</td>
</tr>
<tr>
<td>1991</td>
<td>161,000</td>
<td>60,000</td>
<td>37.3</td>
<td>143,000</td>
</tr>
<tr>
<td>1992</td>
<td>168,000</td>
<td>66,000</td>
<td>39.3</td>
<td>146,000</td>
</tr>
<tr>
<td>1993</td>
<td>170,000</td>
<td>70,000</td>
<td>41.2</td>
<td>149,000</td>
</tr>
<tr>
<td>1994</td>
<td>172,000</td>
<td>72,000</td>
<td>41.8</td>
<td>153,000</td>
</tr>
<tr>
<td>Totals</td>
<td>2,220,000</td>
<td>703,000</td>
<td>34.4</td>
<td>1,961,000</td>
</tr>
</tbody>
</table>
death rates would be most accurately reflected by fatality, not mortality. If fatality-rate reduction was to be accepted as the most important cancer death-rate parameter under these circumstances, then chest x-ray screening has passed a crucial test of efficacy (as reflected by the statistically significant reduction in fatality in the screened group in the Mayo Lung Project).

Figure 2 shows the survival curves in the Mayo Lung Project (considering lung cancer deaths only) and demonstrates superior survival for the experimental group.11 Because incidence is higher in the experimental group and overall lung cancer mortality is the same, these curves are consistent with overdiagnosis. However, existing evidence demonstrates that overdiagnosis is not the result of chest x-ray screening; therefore, the only alternative interpretation of these curves is that the difference between them reflects a benefit from screening. If screening biases (and in particular, overdiagnosis) can be eliminated, then relative survival comparisons would be meaningful, not biased. Accordingly, the 5-year survival rate of 35% in both groups in the Memorial Sloan-Kettering Lung Project7 and 33% in the experimental group in the Mayo Lung Project11 reflect meaningful improvements in outcome when considered in the context of the Mayo Lung Project control group (15% 5-year survival), SEER data (10% 5- to 7-year survival),9 and American Cancer Society cancer statistics (9 to 14% long-term survival, Table 7).

Table 8 demonstrates the potential public health implications had a national policy of annual chest x-ray screening for high-risk individuals been in effect from 1980 to 1994. It uses American Cancer Society estimates (Table 7), and it assumes that the same odds ratio for survival that was observed in the Mayo Lung Project, 2.2, could have been achieved on a national scale. This hypothetical exercise suggests that 315,540 additional patients with lung cancer might have survived their disease since 1980, the year that the recommendation against screening was first promulgated, if a policy of annual chest x-ray screening had been strictly followed.

**CONCLUSIONS**

We believe, based on the analysis presented herein, that existing data from randomized trials are most consistent with the conclusion that periodic chest

![Figure 2](http://journal.publications.chestnet.org/pdfaccess.ashx?url=/data/journals/chest/21716/ on 06/21/2017)

**Table 8—Potential Implications of Policy of Annual CXR* Screening in United States, 1980-1994**

<table>
<thead>
<tr>
<th>Calculation or Source</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total US lung cancer incidence, 1980-1994</td>
<td>Table 7</td>
</tr>
<tr>
<td>Total No. of lung cancer deaths in US, 1980-1994</td>
<td>Table 7</td>
</tr>
<tr>
<td>Total No. of lung cancer survivors in US, 1980-1994</td>
<td>2,220,000-1,961,00</td>
</tr>
<tr>
<td>Fatality rate</td>
<td>1,961,000</td>
</tr>
<tr>
<td>Survival rate</td>
<td>259,000</td>
</tr>
<tr>
<td>Mayo Lung Project survival screened group</td>
<td>Fontana et al11</td>
</tr>
<tr>
<td>Mayo Lung Project survival control group</td>
<td>Fontana et al11</td>
</tr>
<tr>
<td>Mayo Lung Project odds ratio for survival (screen/control)</td>
<td>33</td>
</tr>
<tr>
<td>Projected US survival assuming same odds ratio for survival as observed in Mayo Lung Project</td>
<td>11.7% x 2.2</td>
</tr>
<tr>
<td>Projected total No. of survivors in US (1980-1994) assuming survival rate of 25.7%</td>
<td>2,220,000 x 0.257</td>
</tr>
<tr>
<td>Projected additional survivors of lung cancer in US if policy of annual CXR screening had been followed 1980-1994</td>
<td>570,540-259,000</td>
</tr>
</tbody>
</table>

*CXR=chest x-ray.
x-ray screening is beneficial, as reflected by improvements in stage distribution, resectability, and survival. A case can be made to consider the adoption of periodic chest x-ray screening as standard care for individuals at high risk of lung cancer. Data from the Memorial and Hopkins studies support annual screening as the appropriate interval.

Our conclusion regarding a benefit for periodic chest x-ray screening is at variance with virtually all other previous analyses, which found no benefit for its use in lung cancer screening. Furthermore, because our position is based on interpretation of existing studies rather than on new data, we recognize that many will not accept this conclusion.

The American Cancer Society recently modified its 1980 recommendation against lung cancer screening. It stated (regarding the randomized studies) that it “has reviewed data from these trials and suggests that physicians and high-risk asymptomatic persons determine if chest x-rays are indicated.” However, a recent broad-based, American Cancer Society-sponsored National Conference on Cancer Prevention and Early Detection did not even consider the issue of lung cancer screening. Given the serious public health and fiscal implications of our position in favor of lung cancer screening and the equivocal American Cancer Society position on this issue, we believe that lung cancer screening should be the focus of a national consensus conference to formulate specific guidelines.

Pending such a reassessment of screening guidelines, there is clearly sufficient evidence to justify early detection efforts for lung cancer in the asymptomatic high-risk individual who seeks a cancer checkup from his or her physician. As Fontana et al. have pointed out, “In this day of incomplete medical knowledge and personalized medical practice, it seems reasonable for the individual physician and the individual patient to decide if and when chest radiographs should be obtained. Whether the sum of such individual clinical encounters is equivalent to mass screening is moot.”

ACKNOWLEDGMENT: Data analysis assisted by CDMS data management and analysis system at the Brigham and Women’s Hospital.

REFERENCES
20 McFarlane MJ, Feinstein AR, Wells CK. Clinical features of lung cancers discovered as a postmortem 'surprise.' Chest 1986; 90:520-23
24 Webster E. On the question of cancer induction by small x-ray doses. AJR 1981; 137:647-66
36 Shields PG, Harris CC. Molecular epidemiology and the genetics of environmental cancer. JAMA 1991; 266:681-87