Re: Lung Cancer Mortality in the Mayo Lung Project: Impact of Extended Follow-up

In a recent article, Marcus et al. (1) updated findings from the Mayo Lung Project, extending the follow-up period to 25 years. The authors demonstrated that, even after a long follow-up, there was no difference in lung cancer mortality between the intervention and usual-care arms, although survival for lung cancer diagnosed before 1983 was better in the intervention arm. The authors suggested that overdiagnosis is the most plausible interpretation for the excess of lung cancer incidence observed in the intervention arm. However, alternative explanations for this excess are possible.

One possible explanation for the excess incidence in the intervention arm is an allocation failure, that is, randomization at baseline failed to allocate future lung cancer incidence cases equally into the two arms. An allocation failure, but not overdiagnosis, may explain the results shown in Table 5 of Marcus et al. (1), where there is higher incidence but no survival benefit for patients in the intervention arm with adenocarcinomas diagnosed before 1983. Furthermore, Black editorialized (2) that incidence imbalance could be the result of misdiagnosed occult metastatic adenocarcinomas from other organs, such as primary lung cancers, implying that the number of overdiagnosed cases of lung adenocarcinoma would diminish. For example, suppose that 14 excess deaths due to lung adenocarcinoma were actually due to metastatic adenocarcinoma, then the number of primary lung adenocarcinomas would be 45 and 38 for the intervention and the usual-care arms, respectively, resulting in seven excess cases in incidence in the intervention arm. Although this assumption may be extreme because the number of nonlung cancer deaths was already more in the intervention arm than in the usual-care arm [57 versus 37 cases, respectively; Table 3 in Marcus et al. (1)], it may imply that the majority of incidence excess in the intervention arm came from squamous cell carcinoma: 17 cases of squamous cell carcinoma [Table 5 in Marcus et al. (1)] compared with seven cases of adenocarcinoma. Because more than 70% of squamous cell carcinoma was detected by chest x-ray in the Mayo Lung Project (3) and overdiagnosis due to sputum cytology was not observed in the Memorial Sloan-Kettering and the Johns Hopkins trials (4,5), the excess incidence of squamous cell carcinoma in the study by Marcus et al. (1) could be mainly due to detection by chest x-ray. This contradicts the radiologic findings in Japan that squamous cell carcinomas usually show more rapid growth (6), which means overdiagnosis is less likely than adenocarcinomas. Therefore, when divided into histologic type, overdiagnosis alone could not explain incidence imbalance between the two arms.

It is still unknown whether an excess of lung cancer incidence in the intervention arm continued until the end of the observation period in 1996 because Marcus et al. (1) did not observe lung cancer incidence after 1983. There is a possibility that lung cancer with a relatively long preclinical detectable phase not diagnosed in the usual-care arm before 1983 would offset the difference in lung cancer incidence between the two arms at the end of the observation period in 1996 suggesting that no overdiagnosis exists. Marcus et al. (1) claimed that the incidence excess in the usual-care arm after 1983 is improbable because better survival among participants diagnosed after 1983 in the usual-care arm is unlikely. However, if there turns out to be no excess lung cancer incidence after 1983 in the usual-care arm, implying that lung cancer with a relatively long preclinical detectable phase did not exist in the trial, then lung cancer would consist of two distinct diseases: one with a short preclinical detectable phase and the other with a long preclinical detectable phase and corresponding to the overdiagnosed cases in the intervention arm. Again, these two distinct distributions contradict radiologic and pathologic findings that indicate that lung cancer is distributed in a continuous spectrum from aggressive to indolent characteristics in terms of tumor doubling time and pathologic appearance (6,7).

If the above-stated radiologic and pathologic findings in Japan can be applied to patients in the United States, then the lung cancer incidence imbalance that is observed in the Mayo Lung Project cannot be explained by overdiagnosis alone but may be due to a combination of multiple factors, including, but not restricted to, allocation failure, short observation period, underdiagnosis in the usual-care arm, and overdiagnosis in the intervention arm.

Currently, a disease-specific mortality rate is used as the primary endpoint in randomized controlled trials for evaluating cancer screening. Although there may be no alternatives, the disease-specific mortality rate tends to be affected by the intervention itself and, therefore, needs careful interpretation. It is inappropriate to overinterpret the findings from a single study.

TOMOTAKA SOBUE
TOMIO NAKAYAMA

REFERENCES


Notes

Affiliations of authors: T. Sobue, Cancer Information and Epidemiology Division, National Cancer Center Research Institute, Tokyo, Japan; T. Nakayama, Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases, Japan.

Correspondence to: Tomotaka Sobue, M.D., M.P.H., Cancer Information and Epidemiology Division, National Cancer Center Research Institute, 5-1-1 Tsukiji Chuo-ku, Tokyo 104-0045, Japan (e-mail: tsobue@ncc.go.jp).

The Mayo Lung Project article by Marcus et al. (1) and the accompanying editorial by Black (2) made a valiant effort to underplay the statistically insignificant 13% excess mortality ($P = .09$) in the screened arm. It is shocking that this excess mortality in the screened arm has been interpreted as “No lung cancer mortality benefit evident at the end of the study” (1). A 13% reduction in mortality in the screened arm would have been interpreted as a benefit of intervention that did not reach statistical significance. However, calling the 13% excess mortality no benefit is a tribute to statistical honesty pays to the oncologic optimism of saving lives by screening. Excess deaths in the screened arm in the study by Marcus et al. (1) appear to be increasing with time according to the mortality curve of lung cancer deaths from follow-up years 5–20 [Fig. 1 of Marcus et al. (1)]. Similarly, a Czechoslovakian lung cancer screening study compared chest x-ray with the control arm and also found excess mortality in the screened arm (3). Two other randomized lung cancer screening trials (4), the Johns Hopkins and Memorial-Sloan Kettering studies, have not yet published annual cumulative mortality.

The accompanying editorial by Black goes overboard blaming overdiagnosis for the coexistence of survival benefit and excess mortality in the screened arm. It is most inappropriate to discuss survival in a randomized screening trial because the primary purpose of a randomized screening trial is to overcome the lead time bias. Ignoring the mistake, an important interpretation of the coexistence of survival benefit and excess mortality has been completely overlooked. Blaming overdiagnosis acknowledges that diagnosis itself can affect the natural history of disease, analogous to the impact observation has on the observed. The impact of diagnosis on natural history challenges the paradigm of biological predeterminism and makes way for a new paradigm, referred to as the surgical dissemination/autonomy paradigm. The new paradigm suggests that micrometastases are either disseminated at the time of surgery or that removal of the primary tumor bestows autonomy on pre-existing micrometastases. The surgical dissemination/autonomy paradigm would be ideally tested in a randomized trial comparing surgery with no surgery. However, such a trial would be unethical. Randomized screening trials offer the next best opportunity to test the surgical dissemination/autonomy paradigm. In screening trials for solid tumors, surgery is advanced in the screened arm compared with the control arm. The early years of follow-up after randomization should, therefore, offer an opportunity to compare immediate surgery (screened arm) with the natural progression of the disease for a few years (control arm). Because historically surgery as an event is thought to have little influence on the natural history of cancer, we would predict fewer, if not similar, deaths in the screened arm during the early years of follow-up. If, on the other hand, the surgical dissemination/autonomy paradigm was correct, there should be an excess of deaths in the screened arm during the early years of follow-up. Such excess mortality is indeed evident in the screening trials for lung (1,3), colon (3), and prostate (6) cancers. A meta-analysis of the breast cancer screening trials (7) demonstrated an early excess mortality in the screened arm for the first year of follow-up in women greater than 50 years of age, whereas in women less than 50 years of age, the excess mortality persisted for the first 7 years of follow-up. Thus, published data of screening trials suggests that early intervention apparently has a detrimental effect on the natural history of common solid tumors in the initial years of follow-up.

It would be ideal if a meta-analysis of the annual cumulative mortality by arm of randomization is performed for all lung cancer screening trials. This would allow early detrimental effect and late benefit, if any, to be determined with an adequate sample size.

RAJENDRA A. BADWE

References


Note

Correspondence to: Rajendra A. Badwe, M.D., Department of Surgical Oncology, Tata Memorial Hospital, Parel, Mumbai, India 400012 (e-mail: tnmhcr@vsnl.com)
RESPONSES

We thank Drs. Badwe, Sobue, and Nakayama for their comments regarding our article (1). Dr. Badwe identifies a drawback of cancer screening that is frequently ignored—that screening can result in morbidity and mortality. Badwe suggests that the nonstatistically significant increase in mortality in our study (1) may be due to effects on tumor characteristics. We would like to suggest another possibility—that some individuals with positive screens experience adverse events as a result of diagnostic follow-up. Certain diagnostic evaluation procedures, such as thoracotomy and resection, can result in substantial morbidity and death. Thus, in a cancer screening trial, deaths that occur as the result of positive screen follow-up are (rightly) counted as cancer deaths, because they would not have occurred in the absence of screening.

Drs. Sobue and Nakayama question our conclusion that overdiagnosis is the most plausible explanation for our findings. They provide three alternative explanations. The first alternative explanation is that future lung cancer incidence was not allocated equally among the two arms; in other words, the authors suggest that randomization did not work. We examined this possibility in our article (1) and in an earlier article (2). Furthermore, statistical theory (3) shows that unequal allocation of any risk factor, including those that are either unmeasured or unknown, is an unlikely occurrence, given the large number of study subjects.

The second alternative explanation is predicated on what may be a misinterpretation of the text from Black’s editorial (4). Additionally, Sobue and Nakayama use a definition of overdiagnosis that differs from the one that was used in our article (1) and in Black’s editorial (4). They propose a scenario based on an assumption that they state “may be extreme,” and conclude that overdiagnosis could not alone explain the incidence imbalance, even though they state that “the excess incidence of squamous cell carcinomas…could be mainly due to detection by chest x-ray,” a phrase that in itself describes overdiagnosis as it was defined in our article (1) and in Black’s editorial (4).

The third alternative explanation addresses the issue of catch-up. Sobue and Nakayama argue that it is unlikely that catch-up did not occur because that possibility runs counter to radiologic and pathologic findings. It seems possible that some catch-up has occurred, but whether it has occurred to the extent necessary to result in a similar number of lung cancer cases in each arm is unknown. To address that issue, we have made plans to conduct extended incidence follow-up on the Mayo Lung Project participants. This project is likely to begin in 2001.

PAMELA M. MARCUS
RICHARD M. FAGERSTROM
PHILIP C. PROROK

REFERENCES


NOTES

Affiliation of authors: Biometry Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD.
Correspondence to: Pamela M. Marcus, Ph.D., M.S., National Institutes of Health, 6130 Executive Blvd., MSC 7354, Suite 344, Bethesda, MD 20892–7354 (e-mail: pm145q@nih.gov).

Dr. Badwe states that I have gone overboard in blaming overdiagnosis for the discrepancy in 5-year survival and cumulative lung mortality and completely overlooked an alternative explanation, the “dissemination/autonomy” paradigm. Accordingly, surgical resection of the primary tumor either disseminates micrometastases or bestows autonomy on existing micrometastases. As a radiologist, I have seen several lung cancer patients who had no evidence of distant metastases on preoperative imaging tests but had extensive metastases on tests performed a few months after surgery. Therefore, I believe that the dissemination/autonomy phenomenon does occur in at least some lung cancer patients who undergo surgery.

However, the dissemination/autonomy phenomenon is not inconsistent with overdiagnosis: the diagnosis of lung cancer that would not have otherwise become clinically relevant. To the contrary, the dissemination/autonomy phenomenon will increase cumulative lung cancer mortality substantially only when coupled with overdiagnosis. Without overdiagnosis, the dissemination/autonomy phenomenon can increase the cumulative lung cancer mortality by only the small difference between the cumulative lung cancer mortality and the cumulative incidence of clinically relevant disease, i.e., the probability of dying from minus the probability of dying with clinically relevant lung cancer. Another problem with invoking the dissemination/autonomy phenomenon is that this explanation is not compatible with the Mayo Lung Project data: the impact of confounding and effect modification. J Med Screen 1999;6:47–9.

Nevertheless, I agree with Dr. Badwe’s central premise that screening may actually produce a net harm and not simply fail to produce a statistically significant benefit. For this reason, I believe that any cancer screening program should be proven to be effective with randomized clinical trials before it is recommended outside of the research setting. Furthermore, because the exact cause of death is often multifactorial and difficult to determine, I believe that mortality from all causes should be followed along with disease-specific mortality and carefully considered in the interpretation of the trial results. The examination of all causes of mortality will have the added benefit of putting any observed effect on disease-specific mortality in proper perspective.

WILLIAM C. BLACK

NOTES

Affiliations of author: Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH, and Center for the Evaluative Clinical
Correspondence to: William C. Black, M.D., Department of Radiology, Dartmouth-Hitchcock Medical Center, 1 Medical Center Dr., Lebanon, NH, 03756 (william.black@hitchcock.org).