Editorials represent the opinions of the authors and JAMA and not those of the American Medical Association.

CT Screening for Lung Cancer

Spiraling Into Confusion?

William C. Black, MD
John A. Baron, MD

In this issue of JAMA, Bach et al[1] report their analysis of computed tomographic (CT) screening for lung cancer based on 3 single-arm studies, those from the Instituto Tumori in Milan, Italy, the Mayo Clinic in Rochester, Minn, and the Moffitt Cancer Center in Tampa, Fla. The investigators used a validated lung cancer prediction model to estimate the expected numbers of various lung cancer outcomes among the combined cohort of 3246 participants. To assess the effectiveness of CT screening, they then compared the observed numbers of lung cancer outcomes with the numbers of expected cases. They observed more than a 3-fold increase in the number of new lung cancer cases (144 observed vs 44.5 expected) and a 10-fold increase in lung cancer resections (109 observed vs 10.9 expected). However, there was no decrease in advanced lung cancer cases (42 observed vs 33.4 expected) or in lung cancer deaths (38 observed vs 38.8 expected).

These results follow and are in stark contrast to the recent International Early Action Lung Cancer Program (I-ELCAP),[2] which reported that low-dose CT screening resulted in a 10-year survival of 88% for patients with stage I disease. The investigators argued that CT screening of high-risk individuals could prevent 80% of lung cancer deaths. Subsequently, the lead investigator described the results as "compelling" evidence that CT screening saves lives, and one advocacy group argued that the National Lung Screening Trial,[3] an ongoing multicenter randomized controlled trial of CT screening, should be stopped because it was "outdated" and the effectiveness of CT screening had already been proven.[4]

How is it possible that 2 large studies published within 6 months of each other could lead to such dramatically different conclusions about the effectiveness of CT screening? For a cancer that accounts for more deaths than the 4 next most deadly cancers combined,[5] one group of investigators[1] suggests that CT screening will have no effect on mortality, while the other group[2] suggests that the intervention will reduce mortality by 80%. At least 5 possible explanations for these discordant results deserve some consideration.

One possible explanation is chance. Bach et al[1] observed 38 lung cancer deaths during a median of 3.7 years compared with 38.8 expected deaths during the same period, which results in a relative risk of 1.0 (95% confidence interval, 0.7-1.3). Thus, as the investigators acknowledge, their results are compatible with as much as a 30% reduction in lung cancer mortality. The I-ELCAP investigators[2] observed 75 lung cancer deaths during a median of 3.3 years but did not estimate the expected number of deaths. Therefore, it is not possible to calculate a relative risk from the I-ELCAP results or to construct a 95% confidence interval to determine if it would have been compatible with a 30% reduction in lung cancer mortality.

A second possible explanation for the discordant results is a difference in the populations or screening interventions. Even though the mean age at entry was nearly identical in the 2 studies (60 years vs 61 years), the participants in the study by Bach et al[1] had a stronger smoking history (a mean of 53 pack-years vs a median of 30 pack-years) and were probably at higher lung cancer risk. However, this difference should not have affected the effectiveness of screening expressed as a relative risk. The screening interventions in the 2 studies appear to have been similar, most taking place after the start of the year 2000. In I-ELCAP, the median diameter of screen-detected lung cancer was 13 mm at the prevalence screening and 9 mm at annual screenings.[2] At the Mayo Clinic, the largest contributor to the study by Bach et al in terms of both the number of lung cancers and person-years of observation, the median diameters of screen-detected cancers were 12 mm for the prevalence screening and 10 mm at the annual screening.[6] (The other 2 sites did not report median tumor diameters.) Thus, population characteristics and screening regimens seem unlikely to explain any major differences between the studies.

A third possible explanation that should be considered is the accuracy of the lung cancer prediction model[7] used by Bach et al.[1] This model was derived from a subpopulation of the Carotene and Retinol Efficacy Trial[8] and vali-

See also p 953.

Author Affiliations: Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH (Dr Black); Center for the Evaluative Clinical Sciences, Department of Community and Family Medicine (Dr Black), and Norris Cotton Cancer Center (Dr Baron), Dartmouth Medical School, Hanover, NH.
Corresponding Author: William C. Black, MD, Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756 (william.black@hitchcock.org).

FEB-27-2007 22:52 *77ASSOCIATED PRESS 7 22 P.01

dated against the placebo group of the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study.[9] In the validation, the model underestimated the observed number of lung cancer cases by 11%[9]; the authors hypothesized that this was due to a higher degree of chest radiographic surveillance in the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study than in the Carotene and Retinol Efficacy Trial. If the prediction model used by Bach et al similarly underestimated the expected number of lung cancer cases in the CT screening cohort, the relative risk estimates would have been biased against CT screening. Bach et al acknowledge that the validity of their results depends on the accuracy of the prediction model and appropriately describe their findings as preliminary. There are probably many differences between the Carotene and Retinol Efficacy Trial and Bach et al's CT screening cohort that could have biased the prediction model against or in favor of CT screening. Thus, the prediction model could underlie some of the differences between I-ELCAP and the analysis by Bach et al.

A fourth possible issue that could underlie differences between the 2 studies is the ascertainment of lung cancer deaths. Using the National Death Index and the Italian registry data, Bach et al[1] were able to ascertain vital status and cause of death in more than 99% of screening participants and those for whom vital status could not be reliably determined (n=30) were excluded from the mortality analysis. In the I-ELCAP study,[2] only participants with a known diagnosis of lung cancer, about 1.5% of participants, were followed up annually. Participants without a known diagnosis of lung cancer were followed up for only 1 year after their last CT screening.

Perhaps the best explanation for the contrasting results, however, is the difference in the primary outcome measures of the 2 studies: mortality in the study by Bach et al[1] vs survival in the I-ELCAP study.[2] While these outcome measures are often mistaken to be complementary, prolonged survival in cases need not imply reduced mortality in the population.[10] Case survival can be strongly affected by 3 early detection biases that have no influence on population-based mortality.[11] Lead-time bias occurs when disease is detected earlier but death is not delayed. Because CT screening can detect lung cancer when it is very small, many years of follow-up may be needed to confirm that a death is prevented. Length bias occurs when screening preferentially detects slowly progressive disease. In the Mayo Clinic study,[12] the mean volumetric doubling time of screen-detected lung cancers was 518 days, much longer than that of lung cancers diagnosed before the advent of CT screening, 102 days.[13] In other words, the natural history of small lung cancers detected by CT is unknown. Overdiagnosis occurs when screening detects disease that would not otherwise become clinically evident.[14] Autopsy studies document [redacted] not rare[15] and CT screening is [redacted] more sensitive than [redacted] psy.[16] In a large Japanese study,[17] CT screening detected lung cancer in the same proportion of nonsmokers as smokers, suggesting that many of the screen-detected cancers were not clinically significant. That prolonged survival need not imply reduced mortality also is demonstrated in the study by Bach et al, which found a remarkably high 4-year survival (94%) among clinical stage I lung cancer patients undergoing surgery despite no demonstration of reduction in lung cancer mortality.

Because of the presence of a simulated control group, the measurement of mortality, and the completeness of the outcome ascertainment, the study by Bach et al[1] more directly addresses the population effect of CT screening than does the I-ELCAP study.[2] The study by Bach et al also provides insight into the potential harms of CT screening. A 3-fold increase in lung cancer diagnosis and a 10-fold increase in lung cancer surgery represent substantial psychological and physical burdens. Although the I-ELCAP investigators reported a surgical mortality rate of only 0.5%,[2] Bach et al point out that the average surgical mortality rate across the United States is 5%, and the frequency of serious complications is greater than 20%.[18-21] These potential harms of CT screening mandate that its effectiveness be accurately assessed.

As Bach et al[1] acknowledge, formulation of screening policy should await the rigorous assessment that will be provided by ongoing randomized controlled trials (the National Lung Screening Trial[3] and the NELSON Trial[22]). Randomized controlled trials are the most reliable method for obtaining accurate assessments of the benefits and harms of screening in the underlying population. With this design, differences in outcome can be attributed to the intervention without reliance on highly modeled analyses with problematic assumptions. Although expensive and time-consuming, rigorous trials of cancer screening are far more cost-effective than what might be the alternative—widespread adoption of costly screening interventions that cause more harm than good.

Financial Disclosures: Dr Black reported that he was briefly retained in 2006 as an expert witness by attorneys representing Philip Morris in a class action lawsuit concerning lung cancer screening. His involvement was limited to preparing an affidavit of findings and conclusions set forth in previously published, peer-reviewed medical literature. He did not testify in deposition or at trial. Shortly after the affidavit was submitted, he withdrew from the case and returned the $700 payment he had received due to concern that his service as an expert witness might be misconstrued as support for the tobacco industry. Dr Black reported that he has had no other affiliation whatsoever, including research funding, consultancies, support, or payment of any kind, with the tobacco industry. Dr Baron reported no financial disclosures.

Acknowledgment: We thank H. G. Welch, MD, MPH (Department of Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH; Center for the Evaluative Clinical Sciences, Department of Community and Family Medicine, Dartmouth Medical School, Hanover, NH; and VA Outcomes Group, Veterans Affairs Hospital, White River Junction, Vt), for his careful review of earlier versions of the manuscript and many helpful suggestions. No compensation was given to Dr Welch for his contribution.

REFERENCES

1. Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA*. 2007;297:953-961.
2. Henschke CI, Yankelevitz DF, Libby DM, et al; International Early Lung Cancer Action Program Investigators. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med*. 2006;355:1763-1771.

FEB-27-2007 22:53 *77ASSOCIATED PRESS 7 22 P.02

3. National Cancer Institute Web site. National Lung Screening Trial. http://www.cancer.gov/nlst. Accessed February 8, 2007.
4. Goldberg P. Lung screening advocates say verdict's in, attack NCI randomized trial as "outdated." *Cancer Lett*. 2006;32:1-7.
5. Ries LAG, Harkins D, Krapcho M, et al. National Cancer Institute Web site. SEER cancer statistics review, 1975-2003. http://seer.cancer.gov/csr/1975_2003/. Accessed February 8, 2007.
6. Swensen SJ, Jett JR, Hartman TE, et al. CT screening for lung cancer: five-year prospective experience. *Radiology*. 2005;235:259-265.
7. Bach PB, Kattan MW, Thornquist MD, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*. 2003;95:470-478.
8. Omenn GS, Goodman G, Thornquist M, et al. The beta-carotene and retinol efficacy trial (CARET) for chemoprevention of lung cancer in high risk populations: smokers and asbestos-exposed workers. *Cancer Res*. 1994;54(suppl):2038s-2043s.
9. Cronin KA, Gail MH, Zou Z, Bach PB, Virtamo J, Albanes D. Validation of a model of lung cancer risk prediction among smokers. *J Natl Cancer Inst*. 2006;98:637-640.
10. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA*. 2000;283:2975-2978.
11. Black WC. Randomized clinical trials for cancer screening: rationale and design considerations for imaging tests. *J Clin Oncol*. 2006;24:3252-3260.
12. Lindell RM, Hartman TE, Swensen SJ, et al. Five-year lung cancer screening experience: CT appearance, growth rate, location, and histologic features of 61 lung cancers. *Radiology*. 2007;242:555-562.
13. Geddes DM. The natural history of lung cancer: a review based on rates of tumour growth. *Br J Dis Chest*. 1979;73:1-17.
14. Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst*. 2000;92:1280-1282.
15. Chan CK, Wells CK, McFarlane MJ, et al. More lung cancer but better survival: implications of secular trends in "necropsy surprise" rates. *Chest*. 1989;96:291-296.
16. Dammas S, Patz EF Jr, Goodman PC. Identification of small lung nodules at autopsy: implications for lung cancer screening and overdiagnosis bias. *Lung Cancer*. 2001;33:11-16.
17. Li F, Sone S, Abe H, et al. Low-dose computed tomography screening for lung cancer in a general population: characteristics of cancer in non-smokers versus smokers. *Acad Radiol*. 2003;10:1013-1020.
18. Silvestri GA, Handy J, Lackland D, Corley E, Reed CE. Specialists achieve better outcomes than generalists for lung cancer surgery. *Chest*. 1998;114:675-680.
19. Bach PB, Cramer LD, Schrag D, Downey RJ, Gelfand SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. *N Engl J Med*. 2001;345:181-188.
20. Harpole DH Jr, DeCamp MM Jr, Daley J, et al. Prognostic models of thirty-day mortality and morbidity after major pulmonary resection. *J Thorac Cardiovasc Surg*. 1999;117:969-979.
21. Stephan F, Boucheseiche S, Hollande J, et al. Pulmonary complications following lung resection: a comprehensive analysis of incidence and possible risk factors. *Chest*. 2000;118:1263-1270.
22. van Iersel CA, de Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer*. 2007;120:868-874.