

All-Cause Mortality in Randomized Trials of Cancer Screening

William C. Black, David A. Haggstrom, H. Gilbert Welch

Background: The most widely accepted end point in randomized cancer screening trials is disease-specific mortality. The validity of this end point, however, rests on the assumption that cause of death can be determined accurately. An alternative end point is all-cause mortality, which depends only on the accurate ascertainment of deaths and when they occur. We compared disease-specific and all-cause mortality in published randomized cancer-screening trials to indirectly assess the validity of the disease-specific mortality end point. **Methods:** We examined all 12 published randomized trials of cancer screening for which both end points were available (seven of mammography, three of fecal occult blood detection, and two of chest x-ray screening for lung cancer). For each randomized trial, we subtracted disease-specific mortality observed in the screened group from that observed in the control group and all-cause mortality in the screened group from that in the control group. We then compared the differences in these two mortality measures. **Results:** In five of the 12 trials, differences in the two mortality rates went in opposite directions, suggesting opposite effects of screening. In four of these five trials, disease-specific mortality was lower in the screened group than in the control group, whereas all-cause mortality was the same or higher. In two of the remaining seven trials, the mortality rate differences were in the same direction but their magnitudes were inconsistent; i.e., the difference in all-cause mortality exceeded the disease-specific mortality in the control group. Thus, results of seven of the 12 trials were inconsistent in their direction or magnitude. **Conclusion:** Major inconsistencies were identified in disease-specific and all-cause mortality end points in randomized cancer screening trials. Because all-cause mortality is not affected by bias in classifying the cause of death, it should be examined when interpreting the results of randomized cancer-screening trials. [J Natl Cancer Inst 2002; 94:167-73]

Disease-specific mortality is the most widely accepted end point in randomized clinical trials of screening for cancer (1,2). The validity of this end point, however, rests on the fundamental assumption that the cause of death can be determined accurately. This assumption has been seriously challenged by several studies on the accuracy of death certificates (3-9).

All-cause mortality, in contrast, does not require judgments about the cause of death. Instead, all that this end point requires is an accurate ascertainment of deaths and when they occur. Furthermore, all-cause mortality is a measure that can capture unexpected lethal side effects of medical care. Because of the concern that some cardiac interventions may cause noncardiac deaths (10), for example, there has been a trend toward the use

of all-cause mortality as the primary end point in cardiac drug trials (11-13).

In this article, we review the major randomized trials of cancer screening, point out inconsistencies in their disease-specific and all-cause mortality results, and offer explanations for why these inconsistencies may have occurred.

SUBJECTS AND METHODS

We obtained a list of randomized trials of cancer screening from a published authoritative text devoted to this subject (14). The text cited 16 randomized trials for which disease-specific mortality has been reported. Eight of these trials pertained to screening for breast cancer with mammography, three pertained to screening for colorectal cancer with fecal occult blood testing, and five pertained to screening for lung cancer with chest radiography. In addition, we performed an electronic search on PubMed (National Library of Medicine) by using authors' names and other relevant terms to obtain updates on these trials. We excluded four of these randomized trials because neither all-cause mortality nor data from which all-cause mortality could be calculated were reported [the Stockholm mammographic screening trial (15), the London mass radiology trial (16), the Johns Hopkins study (17), and the Memorial Sloan-Kettering study (18)]. Thus, we found 12 randomized trials of cancer screening for which we could obtain disease-specific and all-cause mortality (19-30): seven of mammography, three of fecal occult blood, and two of chest x-rays for lung cancer (Table 1).

For each of these 12 trials, we used the most recent source that allowed us to determine both disease-specific and all-cause mortality. For consistency, we reported mortality as the number of deaths per 10000 person-years of observation and, for each randomized trial, we used the same denominator for disease-specific and all-cause mortality. For four of the randomized trials, we obtained mortality rates directly from the source (19,27,28,30). For the remaining eight randomized trials, we

Affiliations of authors: W. C. Black, Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH, and Department of Community and Family Medicine, Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, NH; D. A. Haggstrom, Department of Medicine, Dartmouth-Hitchcock Medical Center; H. G. Welch, Department of Medicine, Dartmouth-Hitchcock Medical Center, and Department of Community and Family Medicine, Center for the Evaluative Clinical Sciences, Dartmouth Medical School, and Department of Veterans Affairs Outcomes Group, Department of Veterans Affairs Hospital, White River Junction, VT.

Correspondence to: William C. Black, M.D., Department of Radiology, Dartmouth-Hitchcock Medical Center, 1 Medical Center Dr., Lebanon, NH 03756 (e-mail: william.black@Hitchcock.org).

See "Note" following "References."

© Oxford University Press

Table 1. Twelve randomized trials of cancer screening reporting both disease-specific and all-cause mortality*

Trial (Reference No.)	Disease-specific mortality, No. per 10 000 person-years			All-cause mortality, No. per 10 000 person-years			Inconsistency
	Screened (a)	Control (b)	Screening benefit, b – a (95% CI)	Screened (c)	Control (d)	Screening benefit, d – c (95% CI)	
Mammography							
HIP (19)	5.1	6.5	1.4†	73.7	75.4	1.7†	—
Swedish Two-County (20)	2.6	3.8	1.2 (0.5 to 1.8)	120.4	120.3	–0.1 (–4.4 to 4.1)	Direction
Malmö (21)	3.0	3.0	0.0 (–1.2 to 1.2)	95.4	96.7	1.3 (–5.0 to 7.7)	Direction
Gothenburg (22)	1.3	2.4	1.1 (0.0 to 2.0)	29.6	30.1	0.5 (–3.4 to 4.5)	—
Edinburgh (23)	4.3	5.1	0.8 (–0.7 to 2.4)	80.7	100.8	20.1 (13.3 to 26.9)	Magnitude
Canadian 1 (24)	1.6	1.0	–0.6 (–1.4 to 0.2)	9.0	8.8	–0.2 (–2.2 to 1.9)	—
Canadian 2 (25)	5.0	5.1	0.1 (–1.4 to 1.6)	41.4	38.9	–2.5 (–6.7 to 1.8)	Direction
Fecal occult blood							
Minnesota (26)	5.4	6.6	1.2 (0.1 to 2.4)	183.6	183.6	0.0 (–7.6 to 7.6)	Direction
Nottingham (27)	6.0	7.0	1.0 (0.1 to 2.0)	211.1	209.9	–1.2 (–6.4 to 3.9)	Direction
Funen (28)	6.5	8.2	1.7 (0.3 to 3.2)	221.0	224.0	3.0 (–4.7 to 10.8)	—
Chest x-ray							
Czechoslovakian trial (29)	33.6	24.7	–8.9 (–20.3 to 2.4)	179.2	153.9	–25.3 (–51.6 to 0.9)	Magnitude
Mayo Lung Project (30)	43.9	39.5	–4.4 (–11.0 to 2.1)	324.8	318.5	–6.3 (–24.1 to 11.5)	—

*CI = confidence interval; HIP = Health Insurance Plan.

†Insufficient data were reported to calculate CIs.

calculated the mortality from the reported number of deaths and person-years of observation. [Person-years of observation were reported for four of these randomized trials (21–23,26) but had to be approximated by multiplying reported number of subjects and years of follow-up for the other randomized trials (20,24,25,29)].

For each randomized trial, we subtracted disease-specific mortality observed in the screened group from that observed in the control group. Similarly, we subtracted the all-cause mortality in the two groups. We compared the differences in these two measures of mortality and considered them to be inconsistent if they satisfied one of two conditions. First, if the differences were not of the same sign, then we considered the differences to be inconsistent in direction. Second, if the differences were in the same direction but the difference in all-cause mortality exceeded the disease-specific mortality in the control group, then we considered the differences to be inconsistent in magnitude. We reasoned that screening alone could not explain a deficit in all deaths that exceeded the number of disease-specific deaths, even if screening prevented all disease-specific deaths. Similarly, we reasoned that screening alone would not likely explain an excess in all deaths that exceeded the number of disease-specific deaths. We also calculated the 95% confidence intervals (CIs) around the differences in mortality (31). All *P* values were from two-sided tests and were based on the *Z* test for differences in two proportions.

RESULTS

Perspective

Two observations concerning the mortality rates for 12 screening trials are striking (Fig. 1). First, disease-specific mortality constitutes only a small proportion of all-cause mortality (3%–16% in the control groups). Second, the differences in all-cause mortality within each trial are generally small.

Inconsistencies

Table 1 shows the disease-specific and all-cause mortality in screened and control groups for the 12 trials. Overall, the dif-

ferences in disease-specific mortality were more favorable toward screening than were the differences in all-cause mortality. In five of the 12 trials, the differences were inconsistent in direction. In four of these five trials, disease-specific mortality was lower in the screened group than in the control group, whereas all-cause mortality was the same or higher. Among the seven studies in which the differences were in the same direction, the difference in all-cause mortality exceeded the disease-specific mortality in the control group in two trials. As is evident in Fig. 1, this inconsistency in magnitude was most dramatic for the Edinburgh mammography trial (23). In summary, seven of the 12 trials had results that were inconsistent in either direction or magnitude.

DISCUSSION

Explanations for Inconsistent Direction

Although the goal of screening is to prevent deaths from the target disease, screening may affect mortality in other ways. On the positive side, earlier detection of the target disease could lead to milder treatment and prevent some treatment deaths. In addition, screening could prevent deaths from other diseases that are detected earlier incidentally. On the negative side, screening could lead to deaths from the evaluation of screening test results or from earlier treatment (of the target or other disease) that would not have occurred without screening.

Inconsistency in direction may be partly explained by the inherent ambiguity of disease-specific mortality. Conceptually, this end point should include any cause of death that is modified by screening, including deaths caused by the target disease, by treatment of the target disease, and by the screening process. However, the actual rules used to determine which deaths count for disease-specific mortality are rarely published with trial results. Furthermore, the determination of cause of death is a complex process that is subject to many sources of error. High rates of variation have been demonstrated in the recording of underlying cause of death on death certificates, especially when multiple causes of deaths are involved (3–7).

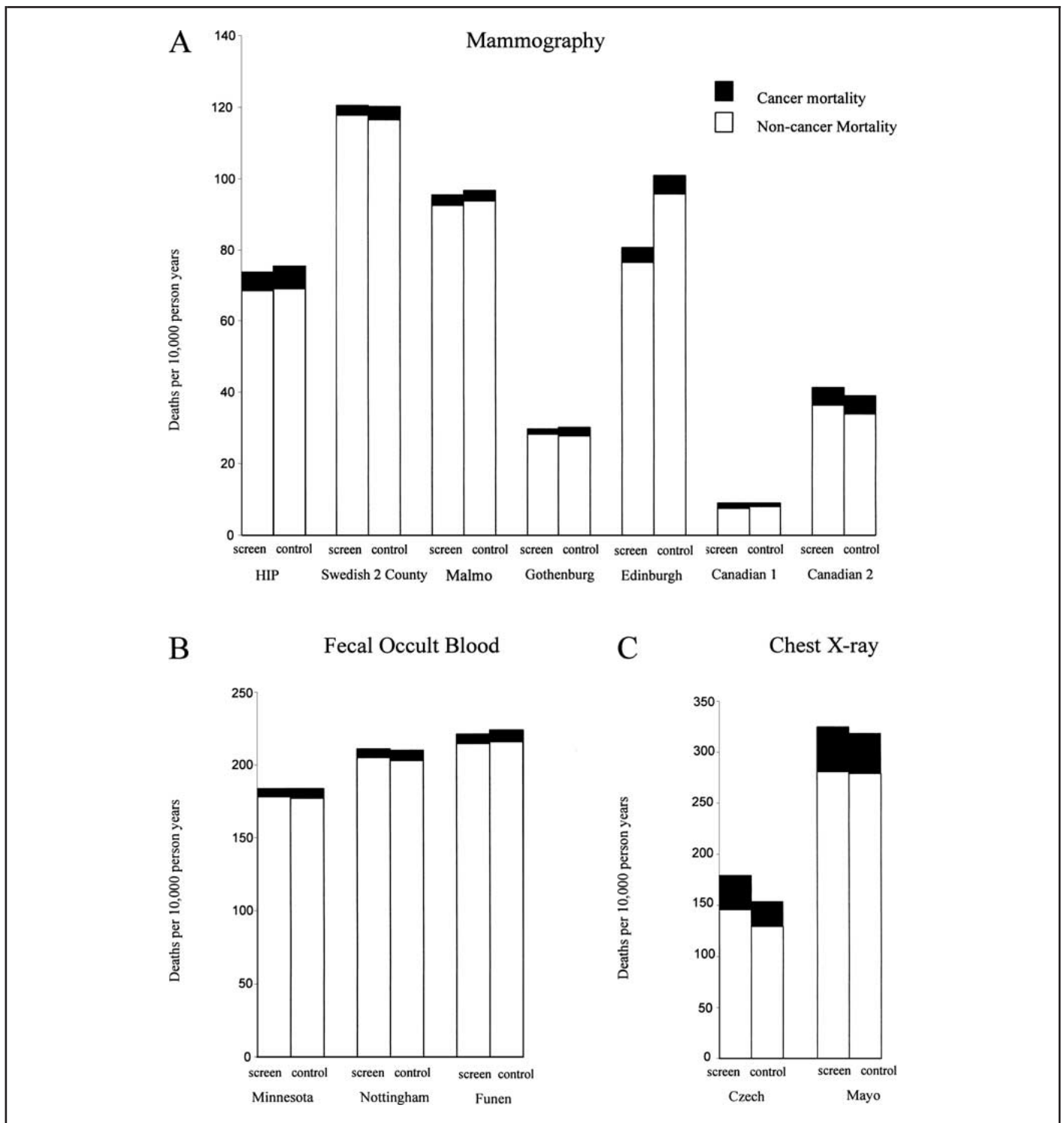


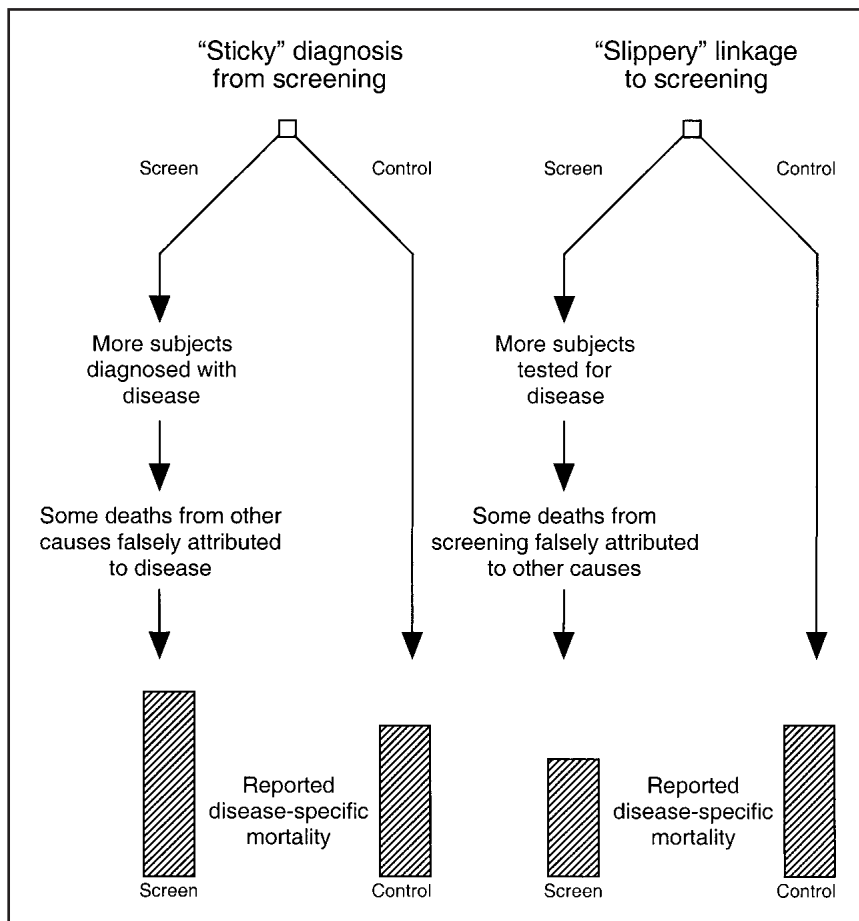
Fig. 1. Mortality rates in randomized trials of screening for cancers of the breast (A), colon (B), and lung (C) (19–30). HIP = Health Insurance Plan.

Two biases may explain much of the observed inconsistency in the direction of disease-specific mortality and all-cause mortality differences (Fig. 2). These biases, which affect only the classification of cause of death, occur independently and have opposite effects on the reporting of disease-specific mortality.

Sticky-diagnosis bias. Because the target cancer in a screening trial is more likely to be diagnosed in the screened group than in the control group, deaths are more likely to be attributed to the target cancer in the screened group. This misattribution

can occur in two different ways. First, a death from another cause in the screened group may be falsely attributed to the target cancer because it had previously been diagnosed in the subject. Second, a death from the target cancer in the control group may be falsely attributed to another cause because the target cancer had not previously been diagnosed in the subject. Although the death-review process in a screening trial can be blinded effectively to randomization, it cannot be blinded completely to the diagnosis or treatment of cancer, which can greatly

Fig. 2. Biases affecting disease-specific mortality. It is assumed that screening has no net effect on mortality; i.e., screening causes the same number of deaths as it prevents. For the sticky-diagnosis bias, deaths from other causes in the screened group are falsely attributed to disease, so that disease-specific mortality is reportedly higher in the screened group. For the slippery-linkage bias, deaths from disease, treatment of disease, or screening process in the screened group are falsely attributed to other causes, so that disease-specific mortality is reportedly lower in the screened group.



influence the subject's subsequent medical care and record. To the extent that the target cancer diagnosis sticks in the screened group (or another diagnosis sticks in the control group), disease-specific mortality will be biased against screening (32,33).

Sticky-diagnosis bias was probably at least partially responsible for the excess lung cancer mortality observed in the screened group of the Mayo Lung Project, and the misclassification was probably most relevant to metastatic adenocarcinoma (30,34). More cases of adenocarcinoma of the lung were diagnosed in the intervention group than in the control group (59 versus 38 cases; $P = .05$), and more deaths were attributed to this cell type in the intervention group (39 versus 25 cases; $P = .10$). In addition, adenocarcinoma was the only lung cancer cell type for which case subjects in the screened group had a shorter median survival than case subjects in the control group, 1.3 versus 1.8 years. Lead-time, length, and overdiagnosis biases should have prolonged survival in the screened group, even if early diagnosis had no beneficial effect. Because the primary site of metastatic adenocarcinoma is often difficult to determine, some deaths from adenocarcinoma of other organs in the intervention group were probably misattributed to lung cancer. In addition, some deaths from adenocarcinoma of the lung in the control group were probably misattributed to adenocarcinoma of other organs or other causes because lung cancer had not been diagnosed previously.

Slippery-linkage bias. Many subjects in the screened group may undergo invasive testing for a suspicious screening result, and many others may be treated for early disease. These interventions may lead to deaths that are difficult to trace back to, or

slip away from, screening. For example, suppose a subject with screen-detected lung cancer has complications of surgery that lead to intensive care. If the subject dies more than 1 month after admission, the death might be falsely attributed to another cause, such as pneumonia. Furthermore, the term disease-specific mortality is too restrictive because it does not imply the inclusion of deaths from screening in individuals without the target disease, such as a fatal hemothorax after a percutaneous needle biopsy for a benign pulmonary nodule. In none of the randomized trials reported in Table 1 is it made clear that such deaths from screening are included in the disease-specific mortality end point or how such deaths would have been identified. To the extent that all these screen-related deaths are misattributed to other causes, disease-specific mortality will be biased in favor of screening.

Slippery-linkage bias may be partly responsible for the discrepancy in the Minnesota Colon Cancer Control Project (26). In this study, there were 1.2 fewer deaths per 10 000 person-years from colon cancer in the screening group than in the control group, but there were 2.1 more deaths per 10 000 person-years from ischemic heart disease in the screening group. Similar discrepancies were observed in the Nottingham trial (27). These findings raise the possibility that the colon cancer screening or subsequent treatment may cause some cardiac deaths that are not properly attributed to the intervention.

Sticky-diagnosis and slippery-linkage biases can both occur in the same trial, and both probably did in the Mayo Lung Project (Table 1). However, that all-cause mortality was higher in the screened group suggests that the slippery-linkage bias was greater than the sticky-diagnosis bias. Thus, more harm than

benefit was probably obscured by the combination of these two biases.

All-cause mortality is not affected by sticky-diagnosis and slippery-linkage biases, because it does not depend on the determination of cause of death. If screening effectiveness (as measured by disease-specific mortality) is obscured because diagnoses are sticking, deaths from other causes will be decreased correspondingly so that all-cause mortality should still reveal a trend toward benefit (albeit not statistically significant). If screening harm is obscured because complications are slipping away from the intervention, deaths from other causes will be increased correspondingly so that all-cause mortality should still reveal a trend toward harm.

Statistical considerations. Because the 95% CIs around the differences in all-cause mortality include zero, all of the inconsistencies in direction in Table 1 could be the result of chance. Of course, that such an important end point has relatively wide CIs is an important observation in itself. Nevertheless, these wide CIs do not imply that chance is the only or even the major cause of the observed inconsistencies in direction, especially when there are plausible alternative explanations.

The problem with conventional statistics in this setting is highlighted by considering a screening intervention that causes far more harm than benefit, yet appears to statistically significantly reduce disease-specific mortality. For example, consider a hypothetical trial of screening in which 100 deaths from the target cancer and 900 deaths from other causes occur in the control group. Suppose that screening prevents 30 deaths from the target cancer but causes 90 deaths that are misattributed to other causes in the screened group. If we assume 10 000 person-years of observation in each arm, then these results could be reported as a statistically significant 30% reduction in disease-specific mortality (relative risk [RR] = 0.70; $P = .02$) and no difference in all-cause mortality (RR = 1.06; $P = .19$).

Explanations for Inconsistent Magnitude

A difference in all-cause mortality that exceeds the disease-specific mortality in the control group is unlikely to be the result of screening, and this difference cannot be explained by bias in the classification of the cause of death. Instead, an inconsistency in magnitude suggests major problems with either randomization or the determination of vital status.

In the Edinburgh trial of breast cancer screening (23) (Table 1), the difference in all-cause mortality between the control group and screened group (20.1 per 10 000 person-years; $P < .001$) was much greater than the breast cancer mortality in the control group (5.1 per 10 000 person-years). These results sug-

gest that healthier subjects were randomly assigned to the screened group. Large discrepancies in socioeconomic status at baseline confirm that there was a major imbalance after randomization, which severely threatens the validity of any comparison between the two groups (35).

Similarly, in the Czechoslovakian trial of lung cancer screening (29) (Table 1), the difference in the all-cause mortality between the control group and the screened group (25.4 per 10 000 person-years; $P = .06$) was greater than the lung cancer mortality in the control group (24.7 per 10 000 person-years). However, in this trial, the two groups were almost identical immediately after randomization. These findings suggest that there was an underreporting of deaths from all causes in the control group, biasing the results against screening.

Table 2 outlines one possible framework for interpreting various combinations of disease-specific and all-cause mortality.

Validity

As we have shown, examination of all-cause mortality in combination with disease-specific mortality can reveal major threats to the validity of a randomized trial, such as flaws in randomization and ascertainment of vital status. In addition, all-cause mortality is unaffected by two biases that affect disease-specific mortality—sticky-diagnosis and slippery-linkage biases. To date, the net effect of these biases appears to have favored screening. In four of the randomized trials where the differences have been inconsistent in direction, the disease-specific mortality has been lower in the screened group. This result suggests that slippery-linkage bias has been a bigger factor than sticky-diagnosis bias and that the harms of screening have been underestimated.

Increasing the rigor of the death-review process might help to reduce the effects of slippery-linkage and sticky-diagnosis biases. However, there are two major limitations inherent in the death-review process. First, it is difficult to devise a search strategy that efficiently identifies all deaths that are plausibly related to screening. In the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial (33), which has the most thorough death-review process ever described, the criteria for death review include a cancer diagnosis or unknown cause of death. This search strategy would not reliably identify fatal complications from an invasive procedure triggered by a false-positive screening test (in a patient who did not have a cancer diagnosis), especially if the subject died after being discharged from the hospital. Second, even if all the deaths could be reviewed, assigning cause of death would still be problematic. For example, if a screened subject had a fatal myocardial infarction

Table 2. Interpretations for various combinations of disease-specific mortality and all-cause mortality

Mortality, screen vs. control	Interpretation	Conclusion
Inconsistent direction		
All-cause lower, disease-specific not lower	Sticky-diagnosis bias: too many deaths attributed to target cancer in the screened group (or too few deaths attributed to target cancer in the control group)	Screening likely effective
All-cause higher, disease-specific not higher	Slippery-linkage bias: too few deaths from work-up and treatment attributed to screening	Screening likely harmful
Inconsistent magnitude	Major flaw in randomization or in ascertainment of vital status	None possible
Consistent		
All-cause lower, disease-specific lower	No net effect of sticky-diagnosis or slippery-linkage bias	Screening effective
All-cause higher, disease-specific higher	No net effect of sticky-diagnosis or slippery-linkage bias	Screening harmful

after an invasive evaluation, it would be difficult to determine whether the death was caused by the evaluation and, thus, attributable to screening.

The main argument for using disease-specific mortality instead of all-cause mortality is that the latter requires far more person-years of observation to generate a statistically significant effect when screening is effective (1). However, all-cause mortality should not be abandoned completely because of a large sample size requirement. Instead, the highest risk populations should be targeted so that the ratio of disease-specific mortality to all-cause mortality is much greater than that in the general population. This would reduce the required person-years of observation. The selection of a high-risk population would also help to avoid the misinterpretation of statistical significance that could make a harmful screening intervention appear to be beneficial or vice versa.

Perspective

All-cause mortality also puts the magnitude of expected benefit from screening into an appropriate perspective for prospective decision making. Promotional efforts at screening usually emphasize death from the target disease, which can lead potential participants to overestimate the impact that screening will have on their probability of dying (36). One might argue that providing information on all-cause mortality would confuse the prospective screenee because the decision to be screened would appear to be a very close call, at least in terms of living or dying. However, this is precisely the case for most forms of screening, and knowing this may reassure the prospective screenee that whatever he or she decides, a major adverse outcome is highly unlikely. This reassurance that the death outcomes are likely to be very similar would be extremely helpful for getting individuals to enroll in randomized trials of screening and to comply with their assignment. In addition, information on all-cause mortality might appropriately redirect the prospective screenee to consider more seriously other interventions that may provide more expected benefit to the individual, such as smoking cessation.

In conclusion, disease-specific mortality may miss important harms (or benefits) of cancer screening because of misclassification in the cause of death. Therefore, this end point should only be interpreted in conjunction with all-cause mortality. In particular, a reduction in disease-specific mortality should not be cited as strong evidence of efficacy when the all-cause mortality is the same or higher in the screened group.

REFERENCES

- (1) Morrison AS. Introduction. In: Screening in chronic disease. 2nd ed. New York (NY): Oxford University Press; 1992. p. 3–20.
- (2) Prorok PC, Kramer BS, Gohagan JK. Screening theory and study design: the basics. In: Kramer BS, Gohagan JK, Prorok PC, editors. Cancer screening: theory and practice. New York (NY): Marcel Dekker; 1999. p. 29–53.
- (3) Hoel DG, Ron E, Carter R, Mabuchi K. Influence of death certificate errors on cancer mortality trends. *J Natl Cancer Inst* 1993;85:1063–8.
- (4) Lee PN. Comparison of autopsy, clinical and death certificate diagnosis with particular reference to lung cancer. A review of the published data. *APMIS Suppl* 1994;45:1–42.
- (5) Messite J, Stellman SD. Accuracy of death certificate completion: the need for formalized physician training. *JAMA* 1996;275:794–6.
- (6) Maudsley G, Williams EM. “Inaccuracy” in death certification—where are we now? *J Public Health Med* 1996;18:59–66.
- (7) Lloyd-Jones DM, Martin DO, Larson MG, Levy D. Accuracy of death certificates for coding coronary heart disease as the cause of death. *Ann Intern Med* 1998;129:1020–6.
- (8) Newschaffer CJ, Otani K, McDonald MK, Penberthy LT. Causes of death in elderly prostate cancer patients and in a comparison nonprostate cancer cohort. *J Natl Cancer Inst* 2000;92:613–21.
- (9) Albertsen P. When is a death from prostate cancer not a death from prostate cancer? [editorial]. *J Natl Cancer Inst* 2000;92:590–1.
- (10) Hulley SB, Walsh JM, Newman TB. Health policy on blood cholesterol. Time to change directions [editorial]. *Circulation* 1992;86:1026–9.
- (11) Pitt B, Poole-Wilson PA, Segal R, Martinez FA, Dickstein K, Camm AJ, et al. Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial—the Losartan Heart Failure Survival Study ELITE II. *Lancet* 2000;355:1582–7.
- (12) Packer M, Coats AJ, Fowler MB, Katus HA, Krum H, Mohacs P, et al. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med* 2001;344:1651–8.
- (13) A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. *N Engl J Med* 2001;344:1659–67.
- (14) Kramer BS, Gohagan JK, Prorok PC, editors. Cancer screening: theory and practice. New York (NY): Marcel Dekker; 1999.
- (15) Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years—update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat* 1997;45:263–70.
- (16) Brett GZ. Earlier diagnosis and survival in lung cancer. *Br Med J* 1969;4:260–2.
- (17) Tockman MS. Lung cancer screening: the John Hopkins study. *Chest* 1986;89(suppl):324.
- (18) Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest* 1984;86:44–53.
- (19) Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: the Health Insurance Plan Project and its sequelae, 1963–1986. Baltimore (MD): John Hopkins University Press; 1988. p. 66–89.
- (20) Tabar L, Fagerberg G, Duffy SW, Day NE. The Swedish Two-County Trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health* 1989;43:107–14.
- (21) Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *BMJ* 1988;297:943–8.
- (22) Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Eriksson O, et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39–49 years at randomization. *Cancer* 1997;80:2091–9.
- (23) Roberts MM, Alexander FE, Anderson TJ, Chetty U, Donnan PT, Forrest P, et al. Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet* 1990;335:241–6.
- (24) Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *CMAJ* 1992;147:1459–76.
- (25) Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50–59 years. *J Natl Cancer Inst* 2000;92:1490–9.
- (26) Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med* 1993;328:1365–71.
- (27) Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 1996;348:1472–7.
- (28) Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996;348:1467–71.
- (29) Kubik A, Parkin DM, Khlat M, Erban J, Polak J, Adamec M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int J Cancer* 1990;45:26–33.
- (30) Marcus PM, Bergstralh EJ, Fagerstrom RM, Williams DE, Fontana R, Taylor WF, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 2000;92:1308–16.

- (31) Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York (NY): John Wiley & Sons; 1981.
- (32) Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer—part II: Cause of death misclassification and the recent rise and fall in prostate cancer mortality. *J Natl Cancer Inst* 1999; 91:1025–32.
- (33) Miller AB, Yurgalevitch S, Weissfeld JL. Death review process in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21(6 Suppl):400S–406S.
- (34) Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening [editorial]. *J Natl Cancer Inst* 2000;92:1280–2.
- (35) Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129–34.
- (36) Black WC, Nease RF Jr, Tosteson AN. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. *J Natl Cancer Inst* 1995;87:720–31.

NOTE

Manuscript received July 2, 2001; revised November 13, 2001; accepted December 4, 2001.