It Is Time to Initiate Another Breast Cancer Screening Trial

Peter Jüni, MD, and Marcel Zwahlen, PhD

In 2009, the U.S. Preventive Services Task Force triggered a heated debate by recommending biennial mammography screening for women aged 50 to 74 years and individualized decisions about mammography for younger women based on context and personal preferences (1). The debate calmed somewhat in 2012 when the Independent U.K. Panel on Breast Cancer Screening concluded that, although associated with some overdiagnosis, screening mammography reduced breast cancer mortality (2). Recent reports have refueled controversy, including a technology assessment issued by the Swiss Medical Board that did not recommend mammography screening regardless of age (3) and the 25-year follow-up of the 2 Canadian National Breast Screening Study trials that showed no advantage of mammography screening over regular clinical breast examination or routine care (4).

Discussions about the benefits and harms of screening commonly focus on the methods and conduct of the trials that tested screening interventions. Trials varied considerably in mechanisms used to allocate women to screening or control groups and approaches chosen for classification of deaths. Six of 11 trials described in a recent review (5) used individual randomization, 4 had a cluster design, and 1 allocated interventions according to birth date. Only 3 reported blind adjudication of deaths.

On the basis of methodological concerns, most experts disregard the findings of 1 of the trials (a cluster trial from Edinburgh), but expert opinion about the remaining trials differ. Opponents of screening point to methodological issues with the Swedish cluster trials that found screening reduced breast cancer mortality considerably, whereas proponents criticize the Canadian trials that found screening did not reduce breast cancer mortality. Here, we reexamine the trials using death from causes other than breast cancer as a marker of proper trial design and conduct.

We posit the following: Mammography provides diagnostic information about the breast, but it does not incidentally detect treatable causes of death in other organs. Therefore, screening mammography should, if effective, reduce breast cancer deaths but not deaths from other causes. Evidence of a benefit of screening on deaths from other causes thus suggests baseline imbalances due to selection bias or the play of chance, or performance bias due to differences in general care that favor the screening group (6). Conversely, mammography should not result in a significant increase in deaths from other causes unless it is severely harmful. Trial results that link mammography to an increased risk for death from other causes suggest imbalances at baseline in favor of the control group or detection bias due to differential misclassification of deaths that favored the screening group on breast cancer deaths but the control group on deaths from other causes (6).

See also:

Related article 861

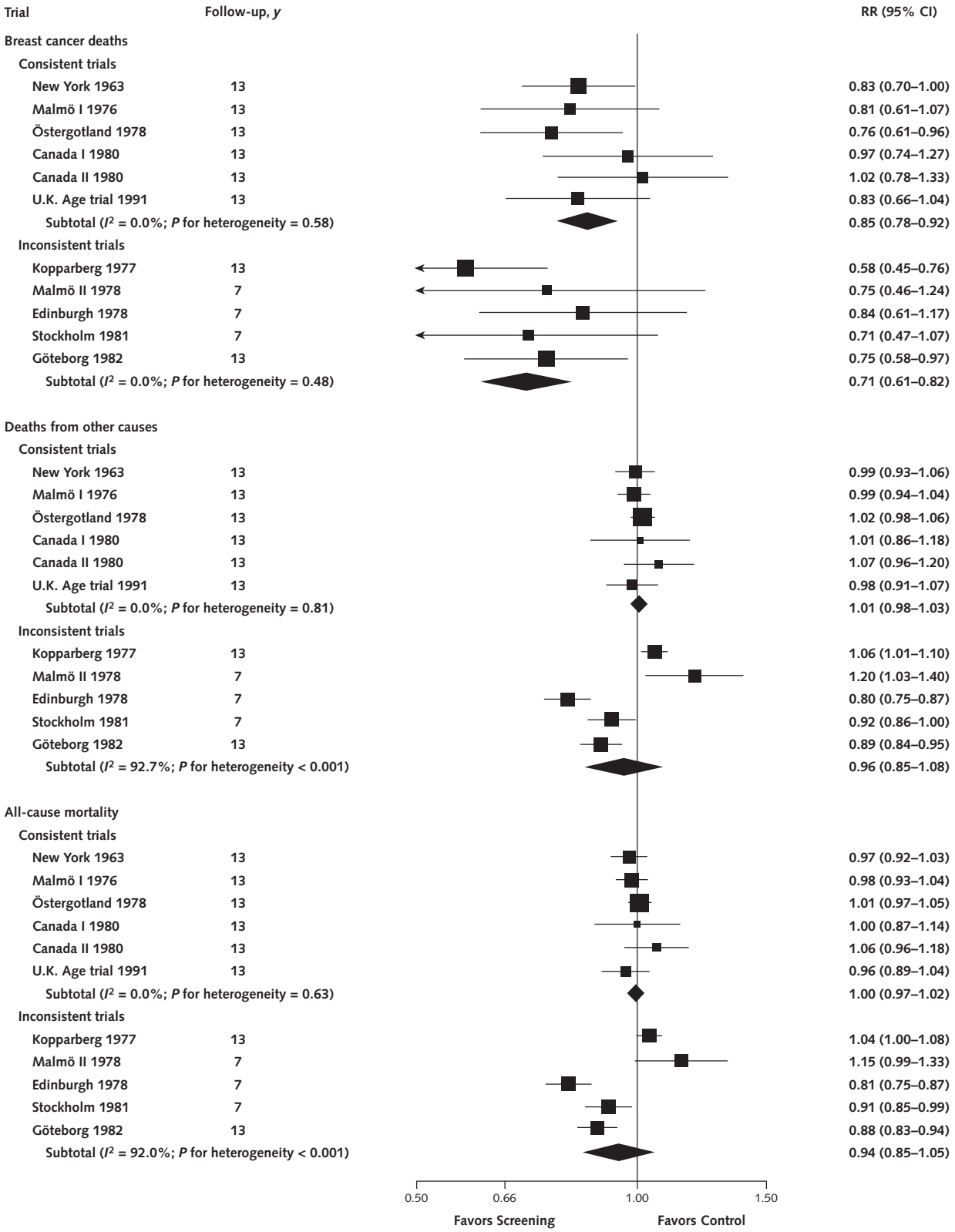
To examine our hypothesis, we plotted the estimated effects of mammography screening on deaths from causes other than breast cancer against the statistical precision of the 11 screening trials (**Appendix Figure**, available at www.annals.org). We used this approach to identify trials in which the cumulative interplay of biases and chance might have rendered results unreliable. Of the 11 trials, 6 were within and 5 were outside the boundaries and showed significant benefit or harm on this outcome. We consider trials with results within the boundaries as "consistent" and those outside the boundaries as "inconsistent" and potentially unreliable. Of the 6 trials within the boundaries, 5 used proper randomization to allocate women individually to mammography or a control group and 3 reported blind adjudication of deaths. Of the 5 trials outside the boundaries, only 1 used individual randomization and none reported blind adjudication.

The **Figure** shows the results of meta-analyses for consistent and inconsistent trials for breast cancer deaths, deaths from other causes, and all-cause mortality after 7 to 13 years of follow-up. Consistent trials suggest that screening results in a 15% reduction of breast cancer deaths and a null result for deaths attributed to other causes. Inconsistent trials show a more pronounced reduction in breast cancer deaths of 29%, whereas the overall pooled relative risk (RR) of deaths from other causes with screening is 0.96, with wide CIs and large between-trial heterogeneity ($I^2 = 93\%$). Inclusion of the inconsistent trials in a pooled analysis probably biases estimates for breast cancer deaths and inflates uncertainty for other types of death.

The ultimate question for any cancer screening program is whether a reduction in cancer mortality really translates into saved lives (7). On average, 10% of the deaths in the control groups of the mammography trials were due to breast cancer (interquartile range, 4% to 14%). A 20% reduction of breast cancer deaths would translate into a 2% reduction of all-cause mortality, which should be detectable given the numbers of women included in the trials: A single trial with 660 000 women and a control group risk of 10% during 15 years of follow-up would have nearly 80% power to detect this effect. Our analysis shows that consistent trials had an RR of 1.00 for all-cause mortality, with a narrow CI ranging from 0.97 to

This article was published online first at www.annals.org on 8 April 2014.

Figure. **Forest plot of random-effects meta-analyses of relative risks for breast cancer deaths, deaths from other causes, and all-cause mortality after 7 to 13 years of follow-up, stratified by consistency of trials.**



Data are from reference 5. RR = relative risk.

1.02 and no heterogeneity (**Figure**). Conversely, inconsistent trials had an imprecise pooled RR of 0.94 and large heterogeneity (I^2 = 92%). Taken together, these analyses do not provide any evidence for an effect of mammography screening on all-cause mortality—the most definitive measure of the net benefit of mammography.

More than 50 years have passed since the first mammography screening trial was initiated. None of the trials were conducted in the era of modern breast cancer treatment. How certain is the magnitude, or even the existence, of a benefit of mammography screening today? Does the currently assumed benefit of screening outweigh the harms of breast cancer overdiagnosis? Were the small reductions in breast cancer deaths afforded by screening in published trials not only diluted but canceled out by deaths from other causes? Should we get accustomed to the possibility that mammography screening does not really save lives? Endless rehashing of data from old trials cannot provide definitive answers to these questions. The only way to know for certain is to initiate a new trial in the era of contemporary screening technologies and breast cancer therapies.

From Institute of Social and Preventive Medicine, University of Bern, Switzerland.

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M14-0569.

Requests for Single Reprints: Peter Jüni, MD, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012, Bern, Switzerland; e-mail, juni@ispm.unibe.ch.

Current author addresses and author contributions are available at www.annals.org.

Ann Intern Med. 2014;160:864-866.

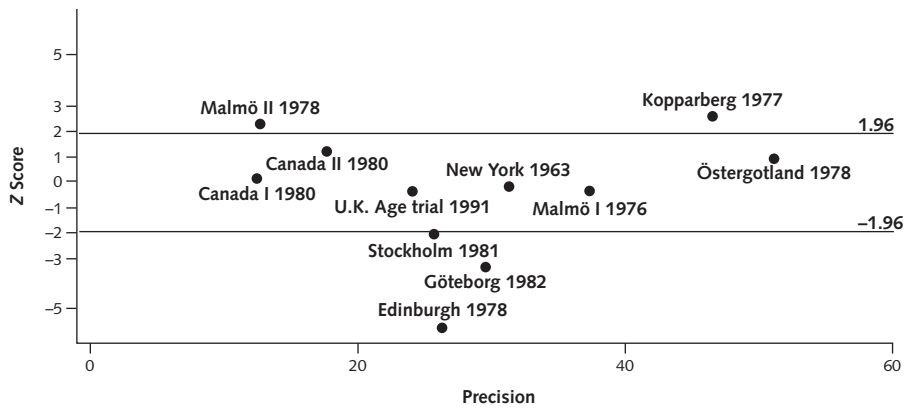
References

1. **U.S. Preventive Services Task Force.** Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009; 151:716-26. [PMID: 19920272]
2. **Independent UK Panel on Breast Cancer Screening.** The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380:1778-86. [PMID: 23117178]
3. **Swiss Medical Board.** Systematic mammography screening. Zollikon, Switzerland: Swiss Medical Board; 2013. Accessed at www.medical-board.ch/fileadmin/docs/public/mb/Fachberichte/2013-12-15_Bericht_Mammographie_Final_Kurzfassung_e.pdf on 3 February 2014.
4. **Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA.** Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ. 2014;348:g366. [PMID: 24519768]
5. **Gøtzsche PC, Jørgensen KJ.** Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2013;6:CD001877. [PMID: 23737396]
6. **Jüni P, Altman DG, Egger M.** Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ. 2001;323:42-6. [PMID: 11440947]
7. **Black WC, Haggstrom DA, Welch HG.** All-cause mortality in randomized trials of cancer screening. J Natl Cancer Inst. 2002;94:167-73. [PMID: 11830606]

Current Author Addresses: Drs. Jüni and Zwahlen: Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012, Bern, Switzerland.

Author Contributions: Conception and design: P. Jüni, M. Zwahlen.
Analysis and interpretation of the data: P. Jüni, M. Zwahlen.
Drafting of the article: P. Jüni.
Critical revision of the article for important intellectual content: P. Jüni, M. Zwahlen.
Final approval of the article: P. Jüni, M. Zwahlen.
Statistical expertise: P. Jüni, M. Zwahlen.
Administrative, technical, or logistic support: P. Jüni.
Collection and assembly of data: P. Jüni, M. Zwahlen.

Appendix Figure. **Modified Galbraith plot of the estimated effects of mammography screening on deaths from causes other than breast cancer against the statistical precision of 11 screening trials.**



The Z score was calculated as ln(RR)/[SE of the ln(RR)]; statistical precision was calculated as 1/[SE of the ln(RR)]. The fixed Z score boundaries at -1.96 and 1.96, represented by the solid lines, divide the plot into areas of significant differences between the screening and control groups ($Z < -1.96$ and $Z > 1.96$, respectively) and nonsignificant differences ($-1.96 < Z < 1.96$). Three trials (Edinburgh 1978, Göteborg 1982, Stockholm 1981) are below the bounds and are associated with a significant benefit of mammography screening on deaths from other causes, whereas 2 others (Malmö II 1978 and Kopparberg 1977) are above the bounds and are associated with a significant harm from mammography screening. If the true RR equals 1, then 1 trial will be outside the boundaries with a probability of 43.1%, 2 trials with 10.2%, and 3 trials with 1.5%. The probability that 5 trials lie outside the boundaries, as is the case, is 0.01%. Data are from reference 5. RR = relative risk.