Responses to the Canadian article published in British Med J

The evaluation of mammography impact on breast cancer mortality in this study poses major challenges that the primary analyses of disease-specific deaths do not overcome. Consequently, the primary mortality analyses which show no benefit of mammography during the screening interval are not as informative about mammography benefit as one might hope.

There are two mortality analyses.

The first looks at the breast cancer death rate restricted to the cases detected during the first five years (the screening period). This is the cumulative death rate in the population but only allows deaths from the cases diagnosed in the first five years. The analysis finds similar death rates on the two arms.

This seems strange because there were there were 666 cases on the mammography arm and 524 cases on the control arm. Thus, if, as the investigators conclude, mammography has no effect, we would actually expect a higher observed cumulative death rate on the mammography arm, unless all of the (142) excess cases are overdiagnosed which is unlikely. Thus, this analysis may bias results against mammography benefit.

The second mortality analysis looks at all breast cancer deaths. But this includes all the cases diagnosed after the screening period. If, after this point, screening behavior evolves similarly in the two arms then the accumulation of a comparable case population in each arm will lead to a dilution of any mortality effect.

In conclusion, both mortality analyses are likely to yield attenuated disease-specific mortality risk ratios even if there is a benefit of mammography.

Competing interests: None declared

Ruth Etzioni, Statistician
Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, P.O. Box 19024, Seattle, WA 98109-1024, USA

The real tragedy is not that an article was reported which describes the results of a deeply flawed trial - physicians can determine for themselves how to interpret these results. The problem is that major media markets picked up the trial are reported the results as a valid challenge to the routine use of mammography. It is not. Most dramatically the New York Times (The Paper of Record) reported a completely unbalanced view of the study, ignored larger, better studies and left the reader with the impression that mammography just does not matter. This is the height of irresponsibility. The BMJ simply reported the results of a trial. It does make one wonder, however, if other journals demurred, realizing that this paper advertises much more than it sells.

Competing interests: None declared

Patrick I. Borgen, Chairman of Surgery;Breast Surgeon
Maimonides Medical Center/Albert Einstein College of Medicine, 4802 Tenth Avenue Brooklyn New York

Computer scientists have an excellent term for nonsense output when input is faulty--GIGO, garbage in, garbage out. This is, unfortunately, what characterizes the Canadian breast cancer screening trial. The cascade of errors in this trial has been eloquently described by others--from substandard mammography (even for the time), inadequate training of mammography interpreters, poor followup of mammographically detected abnormalities, and nonrandomized allocation of women entered into the trial. High numbers of participants and long-term followup of a poorly executed trial cannot be expected to address the deficiencies present at the outset of this trial.

One need only look at Table 1 to realize that something was seriously awry. In no way do the numbers reflect expected (and observed) results when screening is performed properly. Not only was the average size of the cancers detected by screening nearly the same as that of palpable cancers in the control group but there is no evidence of a decrease in average

tumor size over time as one would expect for incident vs. prevalent cancers.

Although the authors claim that randomization was blinded, the significant excess of breast cancers in the screening arm at year 1 coupled with the relatively large average size of the cancers does not favor an explanation of overdiagnosis. And what happened to overdiagnosis in the subsequent 15 years when the number of cancers detected in the screening arm was equal to or less than in the control arm? Overdiagnosis unquestionably occurs in breast cancer screening with mammography but in a study focused on invasive cancers with average tumor sizes of 1.9 cm, the plausibility of significant numbers of 'overdiagnosed' lesions has to be seriously questioned.

The primary lesson we should learn from the Canadian trial is that screening mammography done poorly does not work. Basing decisions about screening on this trial--the only outlier among eight other randomized controlled trials--is tantamount to having the asylum run by the inmates.

Competing interests: None declared

Eva Rubin, Diagnostic radiologist (breast imager)
Montgomery Radiology Associates, 2055 Normandie Drive Ste 108, Montgomery, AL 36111

Dear Editors,

The recent study by Miller et al. (2014) claims to show that annual mammography screening does not reduce breast cancer mortality. An accompanying editorial concludes that the evidence does not support screening women under 60 (Kalager, Adami, and Bretthauer 2014).

Both conclusions are wrong. In fact, this study’s best estimate is that mammography reduced breast cancer mortality by 11.5 deaths per 100,000 women screened. Based on the 95% confidence interval on the effect, their

results are consistent with it being cost-effective to spend $12,000 per person on such a screening program.

The reduction in 25-year cumulative mortality from breast cancer by 11.5 deaths per 100,000 people screened can be found simply by comparing deaths in the treatment group (500 out of 44925, or 1113.0 per 100,000) to those in the control group (505 out of 44910, or 1124.5 per 100,000). Another way of presenting this result is that they find a hazard ratio of 0.99: the death rate of treatment divided by death rate of controls.

Why then do Miller et al. claim that annual mammography screening does not reduce mortality from breast cancer? Because the 95% confidence interval on the effect of screening includes the possibility that it has no effect: the 95% confidence interval on the hazard ratio ranges from 0.88 to 1.12. (A hazard ratio of 1 means no difference in deaths between the two groups.) Nonetheless, the point estimate is that annual screening saves lives, even while the 95% confidence interval indicates substantial uncertainty.

Given this uncertainty, is continued annual mammography screening worth continuing? Put another way, is it cost-effective? A very rough calculation shows that it might well be. (A full cost-effectiveness calculation is certainly worth doing). Suppose the value per life saved was set at $9.1 million (the value used by the U.S. federal government in 2013.) Then, a mammography screening program for 100,000 people that reduced 11.5 deaths—the point estimate in Miller et al. 2014— has a benefit from life saved of $104.7 million. Thus, a screening program with a cost $1047 per person would be cost-effective. Moreover, the 95% confidence interval from Miller et al. also consistent with a reduction of 135 deaths per 100,000 (i.e. a hazard ratio of 0.88). This would entail a benefit from life saved of $1.23 billion, and justify a program cost of $12,300 per person.

While Miller et al. could not reject the hypothesis that annual mammography screening had no effect, they can also not reject the hypothesis that screening saves 135 lives per 100,000 screened, which would justify a screening program substantial costs. We still don’t know for sure—and there are other costs to consider, such as the risk of over

treatment— but this study's best estimate is that annual mammography screening reduces breast cancer mortality.

References

Miller, A. B., C. Wall, C. J. Baines, P. Sun, T. To, and S. A. Narod. 2014. "Twenty Five Year Follow-up for Breast Cancer Incidence and Mortality of the Canadian National Breast Screening Study: Randomised Screening Trial." BMJ 348: g366–g366. doi:10.1136/bmj.g366.

Kalager, M., H.-O. Adami, and M. Bretthauer. 2014. "Too Much Mammography." BMJ 348: g1403–g1403. doi:10.1136/bmj.g1403.

Competing interests: None declared

Keith Ericson, Assistant Professor of Markets, Public Policy, and Law
Boston University School of Management, 595 Commonwealth Ave, Boston MA 02215

Perhaps the authors of this study and the editors of this journal had not completed their training when the first publications relating to the study began to be published in the early 1990's. At that time there was considerable criticism about the randomization of patients in the study. (1-5) The study was actually designed to evaluate the value of mammography in women ages 40 to 49. It was found that there was an excess of patients with advanced cancer found in women aged 40 -49 allocated to the mammography group. Tarone (5) noted that an excess of women with four or more positive nodes in the mammography group would not have arisen by chance if random allocation had been followed. It was subsequently revealed by nurses participating in the study that women with breast complaints were put in the mammography group.

In the current publication related to this study it is reported that 68 percent of the cancers in the mammography group were palpable. How is this

possible? In addition there was little difference in the mean tumor size or in nodal involvement.

In my own practice the 681 invasive cancers detected on mammography average 6 mm in diameter compared to 13 mm for the 640 cancers detected by clinical examination. Nodal involvement in my patients is 19 percent if the cancer is detected on mammography compared to 39 percent if found on clinical examination. Eighty-five percent of my patients with cancers detected on mammography are treated with breast conservation compared to 77 percent of those detected on clinical examination. Ten year disease-free survival for my patients with mammographically detected cancers is 92 percent compared to 82 percent if the cancer was detected on clinical examination. My results are not exceptional.

The authors of the Canadian study, knowing full well that the randomization was flawed, are being unethical in publishing these results without mention of the early criticisms.

1. Mettlin CJ, Smart CR. The Canadian National Breast Screening Study. Cancer 1993; 72 (Suppl): 1461-1465.

2. Conseil d'Evaluation des Technologies de la Santes (CETS). Screening for breast cancer in women 40-49 years. Montreal: CETS; 1993.

3. Kopans DB, Feig SA. The Cannadian National Breast Screening Study: A critical review. Am J Roentgenol 1993; 161: 755-760.

4. Baines CJ. The Canandian National Breast Screening Study: A perspective on criticisms. Ann Intern Med 1994;,120:326-334.

5. Tarone RE. The excess of patients with advance cancer in young women screened with mammography in the Canadian National Breast Screening Study. Cancer 1995; 75: 997-1003.

Competing interests: None declared

Paul I Tartter, Physician
Mount Sinai Roosevelt Division, 425 West 59th Street, 7A, New York, N.Y. 10019

Dear Editor,

It appears that the authors conclusions that annual mammography does not reduce mortality in women age 40-59 beyond that of physical examination or usual care oversteps the limitations of the study. The correct conclusion is that in the 1980s annual clinical breast examinations and training in self breast exam produced nearly equivalent breast cancer specific mortality rates compared to annual mammography screening coupled with annual clinical breast exams and training in self breast exams.

The reasons are the following:

1) All participants in the randomized clinical trial (RCT) had clinical breast exams and were taught breast self examinations at the beginning of the study. The women 40-49 years of age were thereafter randomized to mammography or usual care but the 50-59 year old women were randomized to mammography and annual physical examination or no mammography and annual physical examination. At no point in the RCT was there an arm without either initial clinical breast exam, training in self breast exam, or usual care.

2) A relatively small number of cancers, 32%, were found by mammography alone in the mammography arm of the study. The rest were palpable and were found by either the patient or at the annual physical exam.

3) The mean tumor sizes found in the control arm are not what would be expected in a study with a nonintervention control arm. In our study of mammography detected breast cancer among women age 40-49 we found mean size of mammography detected tumors to be 1.8 cm and the patient/physician detected (palpable) tumor mean size was 3 cm. (1)

4) There were 10% fewer deaths from breast cancer among the mammography detected tumors (19.8%) compared to the palpable tumors

(30.4%) in the mammography arm of the study indicating an advantage from detection of breast cancer at an earlier non-palpable stage.

The author's assertions of overdiagnosis are hard to accept at face value when even the patients with the smallest non-palpable tumors found in the mammography screening arm had a 20% death rate from breast cancer. In our study of breast cancer mortality and treatment by detection method with cases from 1990 to 2008 tracked by our institutional cohort registry the observed death rate from breast cancer among mammography detected cancers was 4% and among patient or physician detected breast cancers was 11%.(1) All breast cancer discovered in this RCT had a hefty breast cancer death rate of 20 to 33% presenting a situation in which it does not appear any of the breast cancer discovered was not going to present harm to the patient during her lifetime. The high rate of breast cancer specific mortality seen in all arms of the study may be due to treatment that has currently been improved upon but was standard of care in the 1980s.

It is not clear if the technology in use in the 1980s when this particular RCT was conducted are comparable to current technology and is particularly questionable given the small difference in mean tumor size observed between arms. Given the small differences in size observed it is doubtful this study has relevance to current technology and the capacity of mammography to be superior to palpation as a method to detect breast cancer at an earlier more treatable stage.

Reference

1) Malmgren JA, Parikh J, Atwood MK, Kaplan HG. Impact of mammography detection on the course of breast cancer in women aged 40-49 years. Radiology 2012;262(3):797-806.

Competing interests: None declared

Judith Malmgren, Epidemiologist
University of Washington, Seattle WA 98177

Dear Editor,

The Canadian National Breast Screening Study has serious sources of distorsion (methodological issues) that have been already mentioned in previous responses to this article. In addition, we have some concerns regarding how the authors presented their findings that can lead to misleading decisions regarding screening.

We wondered why the authors did not stratify the analysis according to age groups. Evidence indicates that there is more harm and less gain for women aged 40-49 years in terms of the number of deaths prevented. Another reason to stratify according to age groups is that they were not comparable in terms of the control arm. Younger women (40-49 years) allocated to he control arm were told to remain under the care of their family doctor whereas older women (50-59 years) were offered annual physical examinations.

The Independent UK Panel on Breast Cancer Screening stated that although there is no single optimum way to estimate overdiagnosis the two most useful estimates are: from the population perspective, the proportion of all cancers ever diagnosed in women invited to screening that are overdiagnosed and from the perspective of a woman invited to screening, the probability that a cancer diagnosed during the screening period represents overdiagnosis.

The authors missed the opportunity to estimate overdiagnosis using the population perspective. That is, the excess of cancer as the proportion of cancers diagnosed over whole follow-up period in women invited for screening (Table 1).

They did not clarify how they estimated overdiagnosis, there is no explanation why they measured the excess cancer cases at 15 years -

instead of 25 years of follow-up- and why they considered as the denominator 484 screening detected cancers -not reproducible data- .

References

1. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380.1778-86.

- table 1_rapid response.pdf

Competing interests: None declared

Carmen Vidal, Breast Cancer Screening Program Coordinator
Montse Garcia
Catalan Institutet of Oncology - IDIBELL, Av. Gran Via 199-203, 08908 Hospitalet de Llobregat (Spain)

The recent publication by Miller et al. [1] indicating no mortality benefits from mammography as part of the Canadian National Breast Screening Study is fraught with bias and the conclusions deeply suspect.

The study ignores pre-invasive cancers which have a high survival rate and are often detectable by mammography. It is reasonable to assume that if they had been included in the analysis then the mammography arm of the trial would have demonstrated higher survival rates than those reported. It is not safe to assume that all pre-invasive cancers are indolent (ie. that they would never go on to harm the patient), especially since the standard model of malignant tumour growth has invasive cancers first developing through a pre-invasive stage.

The study's results contradict the authors' conclusion that mammography is not assisting in saving women's lives. The section of the Results titled "Breast cancer survival" indicates that "The 25 year survival was 70.6% for women with breast cancer detected in the mammography arm and 62.8% for women with cancers diagnosed in the control arm" which the authors

demonstrate to be a statistically significant difference. This demonstrates a real benefit to women surviving breast cancer thanks to receiving mammographic screening. Had pre-invasive cancers been included in this study the difference in 25 year survival is liable to have been even larger. Concluding that mammographic screening provides no benefit with respect to saving women's lives based on the analysis presented [1] is unfounded and dangerous.

[1] A. B. Miller, C. Wall, C. J. Baines, P. Sun, T. To, S. A. Narod, "Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial," British Medical Journal, 2014;348:g366.

Competing interests: None declared

Jacob Levman, Researcher
University of Oxford, Institute of Biomedical Engineering, Parks Road, Oxford, UK, OX1 3PJ

Having been one of the experts called on in 1990 to review the quality of their mammograms I can personally attest to the fact that the quality was poor (1). To save money they used second hand mammography machines. The images were compromised by scatter since they did not employ grids for much of the trial. They failed to fully position the breasts in the machines so that cancers were missed because the technologists were not taught proper positioning, and their radiologists had no specific training in mammographic interpretation.

The CNBSS's own reference physicist wrote:

"..in my work as reference physicist to the NBSS, [I] identified many concerns regarding the quality of mammography carried out in some of the NBSS screening centers. That quality [in the NBSS] was far below state of the art, even for that time (early 1980's). " (2)

In this latest paper (3) the authors gloss over the fact that only 32% of the cancers were detected by mammography alone. This extremely low number is consistent with the poor quality of the mammography. At least two thirds of the cancers should be detected by mammography alone (4). In their accompanying editorial (5) Kalager and Adami admit that " The lack of mortality benefit is also biologically plausible because the mean tumour size was 19 mm in the screening group and 21 mm in the control group....a 2 mm difference." Poor quality mammography does not find breast cancers at a smaller size and earlier stage and would not be expected to reduce deaths.

The documented poor quality of the NBSS mammography is sufficient to explain their results and all of the above disqualifies the CNBSS as a scientific study of mammography screening, but it was even worse than that. In order to be valid, randomized, controlled trials (RCT) require that assignment of the women to the screening group or the unscreened control group is totally random. A fundamental rule for an RCT is that nothing can be known about the participants until they have been randomly assigned so that there is no risk of compromising the random allocation. Furthermore, a system needs to be employed so that the assignment is truly random and cannot be compromised. The CNBSS violated these fundamental rules (6). Every woman first had a clinical breast examination by a trained nurse (or doctor) so that they knew the women who had breast lumps, many of which were cancers, and they knew the women who had large lymph nodes in their axillae indicating advanced cancer. Before assigning the women to be in the group offered screening or the control women they knew who had large incurable cancers. This was a major violation, but it went beyond that. Instead of a random system of assigning the women they used open lists. The study coordinators who were supposed to randomly assign the volunteers, probably with good, but misguided, intentions, could simply skip a line to be certain that the women with lumps and even advanced cancers got assigned to the screening arm to be sure they would get a mammogram. It is indisputable that this happened since there was a statistically significant excess of women with advanced breast cancers who were assigned to the screening arm compared to those assigned to the control arm (7). This guaranteed that there would be more early deaths

among the screened women than the control women and this is what occurred in the NBSS. Shifting women from the control arm to the screening arm would increase the cancers in the screening arm and reduce the cancers in the control arm which would also account for what they claim is "overdiagnosis".

The analysis of the results from the CNBSS have been suspect from the beginning. The principle investigator ignored the allocation failure in his trial and blamed the early excess of cancer deaths among screened women on his, completely unsupportable, theory that cancer cells were being squeezed into the blood leading to early deaths. This had no scientific basis and was just another example of irresponsibility in the analysis of the data from this compromised trial and he finally retracted the nonsense after making front page headlines (6).

The compromise of the CNBSS trial is indisputable. The 5 year survival from breast cancer among women ages 40-49 in Canada in the 1980's was only 75%, yet the control women in the CNBSS, who were supposed to represent the Canadian population at the time, had a greater than 90% five year survival. This could only happen if cancers were shifted from the control arm to the screening arm. The CNBSS is an excellent example of how to corrupt a randomized, controlled trial. Coupling the fundamental compromise of the allocation process with the documented poor quality of the mammography should, long ago, have disqualified the CNBSS as a legitimate trial of screening mammography. Anyone who suggests that it was properly done and its results are valid and should be used to reduce access to screening either does not understand the fundamentals, or has other motives for using its corrupted results.

References:
1 Kopans DB. The Canadian Screening Program: A Different Perspective. AJR 1990;155:748-749.
2. Yaffe MJ. Correction: Canada Study. Letter to the Editor JNCI 1993;85:94
3. 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 doi: 10.1136/bmj.g366
4. Mathis KL, Hoskin TL, Boughey JC, Crownhart BS, Brandt KR, Vachon CM, Grant
CS, Degnim AC. Palpable presentation of breast cancer persists in the era of
screening mammography. J Am Coll Surg. 2010 Mar;210(3):314-83.
Kopans DB, Feig SA. The Canadian National Breast Screening Study: A Critical Review. AJR 1993;161:755-760.
5. Kalager M, Adami H. Bretthauer M. Too much mammography. Long term follow-up does not support screening women under 60. BMJ 2014;348:g1403 doi: 10.1136/bmj.g1403
6. Kopans DB, Feig SA. The Canadian National Breast Screening Study: A Critical Review. AJR 1993;161:755-760.
7. Tarone RE. The Excess of Patients with Advanced Breast Cancers in Young Women Screened with Mammography in the Canadian National Breast Screening Study. Cancer 1995;75:997-1003.
8. Merz B. Author of Canadian Breast Cancer Study retracts warnings. J Natl
Cancer Inst. 1992 Jun 3;84(11):832-4. PubMed PMID: 1593648.

Competing interests: None declared

Daniel B. Kopans, Professor of Radiology
Harvard Medical School, Boston, Massachusetts USA