

Response to: “[Miller’s] Response to: ‘Beyond the mammography debate: a moderate perspective’”

The Editor
Current Oncology
7 January 2016

Scientific progress is made by building upon the findings of other researchers and confirming those findings through repeated experiment. It also comes through learning from one’s mistakes.

Eight major randomized trials of breast cancer screening using mammography have been conducted to test the efficacy of mammography in contributing to reducing breast cancer mortality. In November 2014, most of the world’s leading breast cancer epidemiologists met for 8 days in Lyon, France, at the International Agency for Research on Cancer, part of the World Health Organization, to review all the studies on breast cancer screening. Anthony Miller and I both attended that meeting. The assembled experts concluded, on the weight of the randomized trials and multiple other studies that used more modern mammography than was used in the randomized trials, that there was sufficient evidence that mammography screening reduces breast cancer deaths among women 50–74 years of age and limited evidence among women 40–49 years of age¹. (A resolution for “sufficient evidence” for screening women 45–49 failed by 1 vote.) The Independent U.K. Panel on Breast Cancer Screening had similar findings². Nevertheless, Miller continues to base his position on screening only on his study, the Canadian National Breast Screening Study (CNBSS), ignoring the results of all of the other studies that found benefit when his did not.

In my response to Steven Narod’s defense of the CNBSS³, I pointed to the critical 1993 review of those studies by Boyd *et al.*⁴, which raised multiple concerns regarding the conduct of the CNBSS. Two of the most important were the apparent imbalance in advanced cancers found in the prevalence round of the CNBSS and the poor diagnostic quality of the images as determined by at least 5 external expert breast radiologists who reviewed CNBSS images⁵. In his letter, Miller, while referring to those concerns as specious, then launches into a masterful circumlocution. Responding to my tongue-in-cheek remark about the apparent danger of having been allocated to the mammography arm of his study, he refers to a speculative publication⁶ suggesting (without presenting much in the way of evidence) that

breast surgery should not be performed because it promotes metastases that could lead to death.

Miller also suggests that one should not expect to influence mortality on the basis of cancer detected at the prevalence screen. That makes little sense. In a previously unscreened population, one expects the prevalent cancers to consist of a mix of more advanced cancers that have had an opportunity to develop (and for which the opportunity to avert death is probably lower), but also some earlier cancers whose detection and treatment are more likely to be efficacious. My remark regarding the danger of being in the mammography arm refers to the imbalance in the proportions of the earlier and more advanced cancers between the two trial arms (addressed in more detail shortly), an imbalance that was not seen in any of the other randomized trials.

Miller describes me as a “critical” member of the CNBSS. I should clarify my role there. Initially, he asked me to provide consultation to the study when Irwin Bross, an American biostatistician, suggested that the radiation from all the mammograms used in the CNBSS would be responsible for killing more women than would be saved by the screening. It was not difficult to demonstrate that that argument was wrong. But what I did learn was that the image quality at CNBSS sites was poor. I implemented a quality control program and did my best to advise Miller and Cornelia Baines, the CNBSS deputy director, that it was necessary to improve both the technology used (which in many cases was already obsolete) and also the techniques used by the technologists producing the mammograms and by the radiologists interpreting them. But the mammography in CNBSS was performed at independent facilities, many of them being privately-owned clinics whose choice of technology, decisions to update equipment and techniques, and training of personnel were largely in the hands of the owners, many of whom were unwilling to invest in such improvement. In fact, Miller was able to secure modest funds to assist some of the facilities in improving equipment, but those changes occurred later in the study, and not all facilities were compliant. The structure of the study simply did not allow rigorous standardization of quality, and I expressed that concern publicly in 1993⁷.

Although the quality of the imaging and image interpretation were certainly problematic, the process for allocating women to the study arms is what has elicited

(and should elicit) even greater concern. Miller speaks of the “perfect balance” in the randomization with respect to traditional risk factors. Because most women were properly randomized, that balance is expected, but does not negate the concern about improper allocation of women with palpable abnormalities at the time of randomization. The greatest risk factor for dying of breast cancer is to have entered the study with advanced cancer. An imbalance of only a few women with such cancers is certainly enough to influence the results of the study markedly, and it appears that this indeed is the case. More women with advanced cancers were assigned to the screening group in CNBSS1.

I pointed to that situation in my recent Point-Counterpoint article in this journal⁸, but it might be worthwhile to demonstrate how sensitive a study such as the CNBSS is to bias in allocation. Although 50,430 women were enrolled in the CNBSS1, after the initial follow-up of 8.5 years, there were a relatively small number of breast cancer deaths [$n = 66$: 38 in the screened group (MP) and 28 in the usual care group⁹], an apparent mortality disadvantage of $38 / 28 = 1.36$. In the first (prevalence) screen, several ($n = 24$) poor-prognosis cancers (4 or more positive axillary lymph nodes) were found. That number is not unexpected in a previously unscreened population, but what was surprising is that 19 of those cancers (17 with palpable abnormalities) appeared in the MP arm and only 5 in the usual care arm. Within statistical fluctuation, one would expect equal numbers in the two arms. In his letter, Miller suggests that the imbalance is the result of differences in how nodal involvement was assessed between the study arms (another potential limitation of the study), but there is no evidence that this was the case¹⁰. He also states that because an abnormality “might have been detected on physical examination of the breasts does not necessarily mean that the cancer was palpable.” Of course not. Screening doesn’t directly detect cancers. It detects abnormalities, some of which are diagnosed as cancer. The point is that the almost all of the poor-prognosis cancers in the MP arm had palpable abnormalities and therefore provided no mammography lead time.

Given that many of the observed breast cancer deaths in CNBSS1 would have been attributable to those cancers (presumably Miller could provide this information), even a small imbalance in the number of those cancers would affect the hazard ratio. A rough example is given in Table 1. The top row gives the observations presented in the first report of CNBSS1⁸. In each row, an additional 1 of those cancers has been shifted from the MP arm to the usual care arm until, at the row shown in boldface type, the result that might be expected if the poor-prognosis cancers had been balanced in the allocation of those cancers is seen.

Whatever the reason for the imbalance, the effect on the conclusions of the CNBSS of shifting only 7 women of the 50,430 between study arms is enormous: a 36% mortality disadvantage of screening becomes an 11% mortality reduction! Given the fact that women received clinical breast examination before registration, and thus, that clinic staff were aware of any palpable abnormalities before registration to the trial arm through an open-list process, such a shift is plausible. In the 25-year follow-up publication, Miller *et al.*¹¹ show—by omitting the cancer

TABLE 1 Effect on outcome in the Canadian National Breast Screening Study 1 of shifting women with advanced breast cancer at the prevalence screen between the screening (MP) and usual-care (UC) trial arms

Poor-prognosis cancers in prevalence screen			Breast cancer deaths		
MP	UC	Difference	MP	UC	Ratio
19	5	14	38	28	1.36
18	6	12	37	29	1.28
17	7	10	36	30	1.20
16	8	8	35	31	1.13
15	9	6	34	32	1.06
14	10	4	33	33	1.00
13	11	2	32	34	0.94
12	12	0	31	35	0.89
11	13	-2	30	36	0.83
10	14	-4	29	37	0.78
9	15	-6	28	38	0.74
8	16	-8	27	39	0.69
7	17	-10	26	40	0.65
6	18	-12	25	41	0.61

deaths associated with the prevalence screening round from their analysis—almost the same result. However, in that work they combine data from CNBSS1 and CNBSS2 (women 50–59), and so it is more difficult to interpret the individual results. In any case, to avoid such problems, this type of randomization would not be permitted in modern trials.

I have declared my potential conflicts of interest. Much of my research involves technology and is therefore carried out in collaboration with industry. I also own shares in a company that develops software to measure breast density. Its greatest use will likely be to identify and divert away from mammography screening those women unlikely to benefit from the procedure. After spending more than \$30 million (in today’s Canadian dollars) on CNBSS, Miller has invested more than 25 years in attempting to defend that study, whose flaws are revealed by its own data. I suggest that this is also a conflict of interest that should be declared.

Mammography screening has limitations. They include reduced sensitivity in some women, including those with very dense breasts, recalls of some women without cancer for further imaging and, occasionally, needle biopsy. Certainly, some screen-detected breast cancers, mainly ductal carcinoma *in situ* (and probably some self-detected cancers) are overtreated. But to assert that screening does not reduce breast cancer deaths—not to mention allow for less use of some of the debilitating, aggressive therapies necessary for women with advanced disease—is a fringe opinion at odds with the evidence and global expert opinion. Furthermore, it is an irresponsible message to convey to women and their health care providers.

Martin J. Yaffe PhD CM
Imaging Research, Sunnybrook Research Institute
Departments of Medical Biophysics and
Medical Imaging, University of Toronto
Smarter Imaging Program, Ontario Institute
for Cancer Research
Toronto, Ontario

CONFLICT OF INTEREST DISCLOSURES

I have read and understood *Current Oncology's* policy on disclosing conflicts of interest, and I declare the following interests: My institution receives funding from GE Healthcare for a research collaboration in the areas of digital breast tomosynthesis and contrast-enhanced digital mammography. I am also a founder and shareholder in Volpara Health Technology, a company that produces software for the purposes of quantifying breast density.

REFERENCES

1. Lauby-Secretan B, Scoccianti C, Loomis D, *et al.* on behalf of the International Agency for Research on Cancer Handbook Working Group. Breast-cancer screening—viewpoint of the IARC working group. *N Engl J Med* 2015;372:2353–8.
2. Independent U.K. Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778–86.
3. Narod SA. Counterpoint re: “Mammography screening—sticking to the science.” *Curr Oncol* 2015;22:177.
4. Boyd NF, Jong RA, Yaffe MJ, Trichler D, Lockwood G, Zylak CJ. A critical appraisal of the Canadian National Breast Cancer Screening Study. *Radiology* 1993;189:661–3.
5. Baines CJ, Miller AB, Kopans DB, *et al.* Canadian National Breast Screening Study: assessment of technical quality by external review. *AJR Am J Roentgenol* 1990;155:743–7.
6. Baum M, Demicheli R, Hrushesky W, Retsky M. Does surgery unfavourably perturb the “natural history” of early breast cancer by accelerating the appearance of distant metastases? *Eur J Cancer* 2005;41:508–15.
7. Yaffe MJ. Correction: Canada Study [letter]. *J Natl Cancer Inst* 1993;85:94.
8. Yaffe MJ. Point: Mammography screening—sticking to the science. *Curr Oncol* 2015;22:174–6.
9. 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. [Erratum in: *CMAJ* 1993;148:718]
10. Tabár L, Chen HH, Yen AM, *et al.* Response to Miller *et al.* *Breast J* 2015;21:459–61.
11. 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.