

# The randomized trial of mammography screening that was not—A cautionary tale

Martin J Yaffe<sup>1</sup> , Jean M. Seely<sup>2</sup>, Paula B. Gordon<sup>3</sup> ,  
Shushiela Appavoo<sup>4</sup> and Daniel B. Kopans<sup>5</sup> 

*J Med Screen*  
2022, Vol. 29(1) 7–11  
© The Author(s) 2021



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/09691413211059461  
journals.sagepub.com/home/msc



## Abstract

Two randomized trials were conducted in Canada in the 1980s to test the efficacy of breast cancer screening. Neither of the trials demonstrated benefit. Concerns were raised regarding serious errors in trial design and conduct. Here we describe the conditions that could allow subversion of randomization to occur and the inclusion of many symptomatic women in a screening trial. We examine anomalies in data where the balance would be expected between trial arms. “Open book” randomization and performance of clinical breast examination on all women before allocation to a trial arm allowed women with palpable findings to be mis-randomized into the mammography arm. Multiple indicators raising suspicion of subversion are present including a large excess in poor-prognosis cancers in the mammography trial arm at prevalence screen. Personnel described shifting of women from the control group into the mammography group. There is compelling evidence of subversion of randomization in Canadian National Breast Screening Study. Mis-randomization of even a few women with advanced breast cancer could markedly affect measured screening efficacy. The Canadian National Breast Screening Study trials should not influence breast screening policies.

## Keywords

Breast cancer, randomization, randomized controlled trial, screening

Date received: 1 October 2021; accepted: 19 October 2021

Two breast screening trials began in Canada in 1980 with the intention of answering two questions: (a) Does routine screening of women aged 40–49 with screen-film mammography (M) and clinical breast examination (CBE) confer a mortality reduction from breast cancer compared to usual care (UC)? (b) Does routine screening with screen-film mammography and CBE in women aged 50–59 allow mortality reduction beyond that obtained with CBE alone? These questions were not definitively resolved by the Health Insurance Plan Trial conducted in the US in the 1960s.<sup>1</sup>

In what is now known as the Canadian National Breast Screening Study 1 (CNBSS 1) trial, women aged 40–49 on entry were allocated to either the screening arm (M + CBE) for up to five annual examinations or to “usual care” in the community. The UC group received an initial CBE upon entry to the study and no further intervention. In CNBSS 2, women in their 50 s were allocated to either mammography plus CBE (M + CBE) or CBE alone. In both trials, women received up to five annual examinations.

The CNBSS investigators published 7-year follow-up results in 1992,<sup>2,3</sup> with extended follow-up in 2000, 2002, and 2014.<sup>4–6</sup> In those publications the authors answered “no” to both of the studies’ questions. From the beginning, however, these trials were plagued by concerns and criticisms regarding the randomization process, poor image quality, flaws in the method of recruitment and other factors.<sup>7–12</sup> Now, 41 years after CNBSS began recruiting, in addition to these problems, there are reports of serious protocol violations related to subversion of the randomization that render the

trial results unreliable. How did this happen? Here, we review some aspects of the design and conduct of the CNBSS trials that led to problems with randomization and inclusion of symptomatic women in what were intended to be trials of screening and have resulted in their findings carrying minimal value for informing policy on breast cancer screening.

## The Randomization Process

### Registration method

The two trials were randomized using an “open book” method.<sup>1,6</sup> For each trial, a list was created in which each

<sup>1</sup>Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

<sup>2</sup>Department of Radiology, University of Ottawa, Ottawa, ON, Canada

<sup>3</sup>Department of Radiology, University of British Columbia, Vancouver, BC, Canada

<sup>4</sup>Department of Diagnostic Imaging, University of Alberta, Edmonton, AB, Canada

<sup>5</sup>Department of Radiology, Harvard Medical School, Boston, MA, USA

### Corresponding author:

Martin J. Yaffe, Rm S657, 2075 Bayview Avenue, Toronto, ON M4N3M5, Canada.

Email: martin.yaffe@sri.utoronto.ca

row represented an identifiable pre-randomized allocation to one trial arm. If participants were registered line by line in the order in which they presented at the study site, the trial would have been properly randomized; however, the open-book system allowed for possible subversion. A participant destined for one trial arm could be entered into the other by skipping the appropriate number of rows. Subsequent participants would then be registered in the blank slots and, after the fact, no sign of subversion would exist

### *Clinical breast examination*

Except at one of the 15 study sites, all women received a CBE by a trained nurse examiner (in Quebec, physicians performed the exam) *before* being registered in the CNBSS.<sup>8</sup> This was not the order specified in the protocol, but for reasons unknown was a practice adopted at 14 of the sites. The examiner would have knowledge of any clinical findings at the time of registration. It was possible, given the registration method, that a study nurse could influence the allocation of a participant to a particular trial arm. A likely possible scenario would be in the case where there was a clinical finding and the nurse, with good intentions, but not appreciating the critical importance of randomization, could have urged that a woman be entered into the M + CBE arm of the trial so that she would receive a mammogram immediately.

Where is the evidence that such events occurred in the CNBSS? It is seen in the first publication of CNBSS 1 (women in their 40 s) results in 1992.<sup>1</sup> Of 24 poor prognosis cancers (four or more positive axillary lymph nodes) found at the prevalence screen, 19 occurred in women registered in the M + CBE arm and only five in the UC arm. Of these 19, 17 were palpable. Boyd et al. in a published critique of the CNBSS calculated that the probability of such imbalance occurring randomly was only 0.0033.<sup>7</sup> They pointed to the open-book registration and the pre-examination as serious concerns.

There were other data that supported this hypothesis. Screening mammography typically finds many cancers before they can be detected clinically. In CNBSS 1 only 32% of the cancers were detected by mammography alone. This is consistent with the increased prevalence of more advanced, clinically evident cancers in the M + CBE arm, causing the fraction of mammographically alone detected cancers to drop.

At one of the screening sites (Winnipeg), an analysis was performed correlating independent previous public health insurance claim data with an assignment to trial arms. Of nine CNBSS 1 and CNBSS 2 participants who had prior health claims for breast cancer, eight were registered in the mammography arms. In CNBSS 1, four of four such women were assigned to the mammography arm.<sup>13</sup>

Why was not this problem discovered in 1992? In fact, it was. Prompted by concerns voiced by Kopans,<sup>8</sup> Boyd et al.<sup>7</sup> and later Tarone,<sup>11</sup> in 1995 an external investigation into possible irregularities with randomization was launched by the trial sponsor, the National Cancer Institute of Canada, and the Faculty of Medicine of the University of Toronto. The external investigators, Drs John Bailar and Brian MacMahon, were

asked to review the randomization procedures that had been used in the CNBSS.

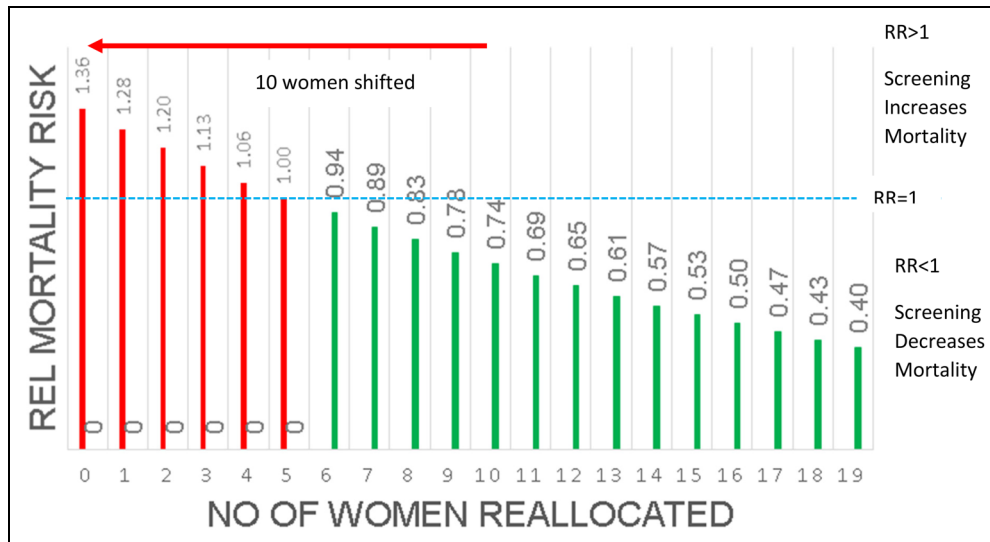
In their report,<sup>14</sup> Bailar and MacMahon stated that “To avoid subversion of randomization of this type, it is current practice to conceal the allocation from both the study subject and the person doing the randomization until or after the commitment of the subject to a particular arm of the study.” But they confirmed that in the CNBSS, formal allocation occurred after the clinical examination and that “the nurses (and probably also the coordinators) were aware of the findings of the clinical examination when the allocation was made.”

Bailar and MacMahon, with the help of forensic document experts (from KPMG Investigation and Securities Inc., Toronto), limited their investigation to looking for erasures/alterations in the allocation books. Alterations were found on 101 lines allocated to mammography. The investigators acknowledged that they visited only 3 of the 15 study sites, and that they “did not fill the terms of reference given to us... we did not interview any of the field staff of the study, even though a few are still working in the participating centres.” This was a critical omission; since the allocation of participants was on open lists, any compromise in the process would have been undetectable. Dr Kopans wrote to Drs. Bailar and MacMahon in 1995, prior to the review, “emphasizing the importance of interviewing the individuals involved in random allocation in the National Breast Screening Study (NBSS). The most direct way to find out whether the process was compromised would be to ask those involved in the allocation and to provide them with anonymity and protection from retribution.” This was not done. Bailar and MacMahon also acknowledged that a coordinator from one of the centers was dismissed because she had assigned her friends to the mammography arm.

Even though they did not interview the staff, they “noted several ways in which the randomization in the NBSS study could have been subverted.” They acknowledged that they could not “exclude the possibility that one or the other of these methods of subversion (or other methods we are unaware of) were used,” but argued that even if subversion had occurred, it “would have affected very few individuals.”<sup>9</sup>

In a simple calculation, one of us (MJY)<sup>15</sup> demonstrated how even a small imbalance in the assignment of women who entered the study with advanced cancers would definitively shift the study results away from a demonstration of mortality reduction. It is seen (Figure 1) that systematically shifting only 10 women with advanced cancer in CNBSS 1 from the UC arm to the M + CBE arm could have the effect of changing a 26% mortality reduction into the appearance of a 36% increase in mortality.

In 1995, Tarone, in response to concerns about randomization in CNBSS, suggested that the data be analyzed excluding the deaths that occurred from cancers detected in the prevalence round.<sup>8</sup> For two decades the CNBSS investigators resisted performing this analysis. Finally, they did such an analysis in 2014; however, unfortunately, and inexplicably, the study investigators chose to pool the data from the two different studies (CNBSS 1 and 2).<sup>6</sup> This is without justification as the studies actually had different designs based on two different



**Figure 1.** The effect of shifting the allocation of women with advanced breast cancers, destined to die, from the usual care arm to the mammography arm of Canadian National Breast Screening Study (CNBSS) I. The red bar indicates that shifting only 10 of these women would cause a mortality reduction benefit of 26% to appear instead as an increased risk of 36%.

hypotheses. Nevertheless, the analysis showed that by eliminating that first screening round the hazard ratio associated with screening fell 15% points from 1.05 (95% confidence interval (CI): 0.69–1.16) to 0.9 (95% CI: 0.85–1.3).<sup>6</sup> Without speculating on the exact nature of the cause of that drop, the change is compelling evidence of a randomization problem. Furthermore, the authors reported that the breast cancer mortality hazard ratio associated with the mammography arm was 1.47 for cancers detected in the prevalence screen, compared with 0.9 in subsequent screens,<sup>6</sup> and that the 7-year survival rate for women in the M + CBE arm was 88.7% compared to 91.2% in the UC arm,<sup>1</sup> further supporting the argument that the initial randomization was not balanced, with the imbalance conferring a bias against observing a benefit from screening.

Bailar and MacMahon concluded that they could find no credible evidence that compromise of the randomization had occurred and they came up with no explanation for the imbalance in allocation. Miller has argued that the fact that overall demographic measures in CNBSS were well balanced between the study arms is proof of valid randomization.<sup>2,3,16</sup> It is, important, however, to realize that shifting even 100 or more women of the 50,430 women in CNBSS I would have no discernible effect on that balance: the imbalance in the small but crucial numbers of *advanced cancers* could markedly alter the comparison of deaths due to breast cancer while making no discernible difference in the risk factor profiles in the study population as a whole.

This left strong suspicions regarding major protocol violations in CNBSS,<sup>17,18</sup> however, the evidence for subversion of randomization remained circumstantial for three decades.

In March 2021, one of us (DBK) presented a virtual invited lecture at the University of Toronto about the CNBSS, mentioning the concern about randomization. After the lecture, a mammography technologist who worked at the St Michael's Hospital site of CNBSS came forward, confirming the

long-held suspicions of compromised randomization.<sup>19</sup> She reported that some women with breast lumps were deliberately assigned to the mammogram arm of the study, rather than blindly randomized. Supportive of this testimony was a very high screen-1 cancer detection rate in the mammography arm at St Michael's site, much higher than at other sites and 12 times higher than that for the CBE arm, which had a very low cancer detection rate. Similarly, the benign-to-malignant ratio at biopsy at this site was low (1.8:1) compared to other sites, suggesting a preponderance of larger, easily detected cancers (unpublished data).

Subsequently, another eyewitness, a radiologist at a separate screening site in a different city, came forward with information about serious protocol violations at that site.<sup>19</sup>

## Recruitment of Symptomatic Women into the Trials

The CNBSS trials were intended to be studies of screening, the detection of breast cancer in asymptomatic women. Women with symptoms are unlikely to benefit from screening and require diagnostic workup, including imaging.

Women who entered the CNBSS trials were volunteers. Mammography was not widely performed as a screening procedure at that time and many of the women who entered the CNBSS were likely to have done so because they had a 50% chance of receiving yearly mammography exams. It is not unreasonable to assume that they may have been concerned about the possibility that they had breast cancer and were, therefore, interested in having a mammogram.

At some screening centers, however, where the study recruitment was lower than desired, women were recruited from the practices of breast surgeons. The likelihood of these women being asymptomatic would have been very low.

In fact, through discussions with individuals who had worked in a professional capacity at CNBSS sites, 17 of these people at seven sites stated that they were aware of women with symptoms being recruited.<sup>18,20</sup> Some of these women undoubtedly would have had breast cancer, detectable either at entry or in subsequent years, and by contributing to deaths in both arms of the study these symptomatic cancers would have added a systematic bias to deaths in both arms, diluting the ability to detect any beneficial effect of mammography screening. This might in part explain the high fraction of palpable cancers found at the initial screen in CNBSS 1 (58.1%) and CNBSS 2 (54.2%).<sup>2,3</sup>

This effect would occur in addition to that associated with subversion of randomization, further reducing the measurement of any real benefit from screening.

## Discussion

The local, open-book method of randomization in CNBSS and the fact that all women received CBE before being formally given their allocation in the trial registry created conditions where it was not difficult to subvert random assignment to the arms of the CNBSS trials, and this could be done without detectable signs such as erasures. The initial CBE also provided motivation for the examiner to direct women with palpable findings into the arm where they would receive mammography immediately. The systematic shifting of CNBSS participants between study arms at entry could have implications beyond deaths due to cancers detected in the prevalence screen. The presence of palpable findings in women at entry is often associated with atypical hyperplasia, which is strongly predictive of an increased risk of cancer that could be detected in subsequent screens or appear as interval cancers, some of which would result in deaths associated with detection at these later time points.<sup>21,22</sup>

Given the documented witness statements,<sup>19</sup> there are key questions remaining that could explain the results of the CNBSS. How many women with more advanced cancers were shifted from their original random allocation in the CNBSS? On an individual basis, how many of the cancer deaths in each study arm were associated with the presence of advanced cancers in that arm at study entry? Similarly, how many deaths in each arm were associated with the presence of palpable signs at study entry as opposed to cancers that were detected by mammography alone? How much variability was there among study sites in the imbalance of advanced cancers between study arms? How much variability was there among study sites in the probability of dying of breast cancer between the two study arms? These are questions that can only be answered by the CNBSS investigators.

This is not simply an issue for the CNBSS. It is a salutary lesson that random allocation in a clinical trial should not be predictable for the recruiting staff and should be unalterable after the fact. One would hope that this lesson has already been learned, but it does no harm to repeat it. This is also a lesson for those who conduct systematic reviews. For many years, concerns about the potential subversion of randomization in these trials have been expressed in the public domain,

yet many systematic reviews dismissed these, taking the demographic similarity of the study arms as a whole (irrelevant to the issue of systematic allocation of the small numbers of trial participants with existing palpable disease, as noted above) as reassurance that the randomization was adequate.<sup>23,24</sup>

## Conclusion

In addition to other serious flaws that have been documented elsewhere,<sup>6–10</sup> there is now direct eyewitness evidence available of subversion of randomization of the CNBSS and the systematic inclusion of symptomatic individuals. CNBSS 1 and 2 were not reliably randomized controlled trials, nor were they truly trials of screening. Their results should no longer be used in meta-analyses of screening nor to inform policies on breast cancer screening. Future trials should avoid the weaknesses in design methodology and should contain effective measures to prevent the flaws in execution that were found in CNBSS 1 and 2.


## Declaration of conflicting interests


The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Research collaboration between Dr Yaffe's institution, Sunnybrook Research Institute, and GE Healthcare on breast tomosynthesis and contrast-enhanced mammography. Yaffe holds shares in Volpara Health Technologies and is the principal of Mammographic Physics Inc., a company that provides consulting on image quality and radiation safety issues in breast cancer imaging. Dr. Gordon holds shares in Volpara Health Technologies. Dr. Kopans is an advisor to DART Imaging which is developing digital breast tomosynthesis devices for China.


## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Martin J Yaffe  <https://orcid.org/0000-0002-7227-9915>

Paula B. Gordon  <https://orcid.org/0000-0002-2477-1278>

Daniel B. Kopans  <https://orcid.org/0000-0003-0741-6829>

## References

- Shapiro S, Strax P and Venet V. Periodic breast cancer screening in reducing mortality from breast cancer. *JAMA* 1971; 215: 1777–1785.
- Miller AB, Baines CJ, To T, et al. Canadian National breast screening study: I. Breast cancer detection and death rates among women aged 40 to 49 years. *CMAJ* 1992 Nov 15; 147: 1459–1476.
- Miller AB, Baines CJ, To T, et al. Canadian National breast screening study. II. Breast cancer detection and death rates among women aged 50 to 59 years. *Can Med Assoc J* 1992; 147: 1477–1488.
- Miller AB, To T, Baines CJ, et al. The Canadian national breast screening study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002; 137: 305–312.
- Miller AB, To T, Baines CJ, et al. Canadian National breast screening study-2: 13-year results of a randomized trial in women age 50–59 years. *J Natl Cancer Inst* 2000; 92: 1490–1499.
- Miller AB, Wall C, Baines CJ, et al. Twenty-five-year follow-up for breast cancer incidence and mortality of the Canadian national breast screening study: randomised screening trial. *Br Med J* 2014 Feb 11; 348: g366.
- Boyd NF, Jong RA, Yaffe MJ, et al. A critical appraisal of the Canadian national breast cancer screening study. *Radiology* 1993 Dec; 189: 661–663.
- Kopans DB. The Canadian screening program: a different perspective. *AJR Am J Roentgenol* 1990; 155: 748–749.
- Kopans DB and Feig SA. The Canadian national breast screening study: a critical review. *AJR Am J Roentgenol* 1993; 161: 755–760.

10. Burhenne LJ and Burhenne HJ. The Canadian national breast screening study: a Canadian critique. *AJR Am J Roentgenol* 1993 Oct; 161: 761–763.
11. 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.
12. Kopans DB. Major failings of trial procedures and quality of screening fatally compromise the results of the Canadian national breast screening studies. *J Med Screen* 2021 Jun; 28: 59–62.. Epub 2021 Jan 17.
13. Cohen MM, Kaufert PA, MacWilliam L, et al. Using an alternative data source to examine randomization in the Canadian national breast screening study. *J Clin Epidemiol* 1996; 49: 1039–1044. [PMID: 8780614].
14. Bailar JC3rd and MacMahon B. Randomization in the Canadian national breast screening study: a review for evidence of subversion. *CMAJ* 1997; 156: 193–199.
15. Yaffe MJ. Letter to the editor. Response to: [Miller's] response to: beyond the mammography debate: a modern perspective. *Curr Onc* 2016 Jun; 23: e324–e326.. Epub 2016 Jun.
16. Miller AB. Final results of the UK Age trial on breast cancer screening age. *Lancet Oncol* 2020; 21: 1125. (letter).
17. Kopans DB. NBSS: opportunity to compromise the process. *Can Med Assoc J* 1997; 157: 247.
18. Boyd NF. The review of randomization in the Canadian national breast screening study. Is the debate over? *CMAJ* 1997; 156: 207–209.
19. <https://dataverse.scholarsportal.info/dataset.xhtml?persistentId=doi:10.5683/SP3/2DEY36> (accessed 30 October 2021).
20. Seely JM, Eby PR, Gordon PB, et al. Errors in conduct of the CNBSS trials of breast cancer screening observed by research personnel. *J. Breast Imaging*, in press.
21. Ernster VL. The epidemiology of benign breast disease. *Epidemiol Rev* 1981; 3: 184–202.
22. Dupont WD and Page DL. Risk factors for breast cancer in women with proliferative breast disease. *NEJM* 1985; 312: 146–151.
23. Nelson HD, Cantor A, Humphrey L, , et al. *Screening for breast cancer: a systematic review to update the 2009 U.S. Preventive services task force recommendation. Evidence synthesis No. 124. AHRQ publication No. 14-05201-EF-1*. Rockville, MD: Agency for Healthcare Research and Quality, 2016.
24. Moher D, Little J, Barbeau P, et al. Breast cancer screening: Part A. An evidence report to inform an update of the Canadian Task Force on Preventive Health Care 2011 Guideline. 2017. [https://canadiantaskforce.ca/wp-content/uploads/2019/02/Systematic-Review-Evidence-Report\\_v2\\_FINAL.pdf](https://canadiantaskforce.ca/wp-content/uploads/2019/02/Systematic-Review-Evidence-Report_v2_FINAL.pdf) (accessed 30 September 2021).