

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/277410788>

Conclusions for mammography screening after 25-year follow-up of the Canadian National Breast Cancer Screening Study (CNBSS)

ARTICLE *in* EUROPEAN RADIOLOGY · MAY 2015

Impact Factor: 4.34 · DOI: 10.1007/s00330-015-3849-2 · Source: PubMed

DOWNLOADS

25

VIEWS

62

5 AUTHORS, INCLUDING:



Sylvia H Heywang-Köbrunner

Referenzzentrum Mammographie München

200 PUBLICATIONS 3,605 CITATIONS

SEE PROFILE



Maria R Noftz

Universität zu Lübeck

5 PUBLICATIONS 36 CITATIONS

SEE PROFILE



Alexander Katalinic

Universität zu Lübeck

350 PUBLICATIONS 4,538 CITATIONS

SEE PROFILE

Conclusions for mammography screening after 25-year follow-up of the Canadian National Breast Cancer Screening Study (CNBSS)

S. H. Heywang-Köbrunner¹ · I. Schreer² · A. Hacker¹ · M. R. Nofzt³ · A. Katalinic³

Received: 28 January 2015 / Revised: 8 April 2015 / Accepted: 13 May 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Twenty-five-year follow-up data of the Canadian National Breast Cancer Screening Study (CNBSS) indicated no mortality reduction. What conclusions should be drawn? After conducting a systematic literature search and narrative analysis, we wish to recapitulate important details of this study, which may have been neglected: Sixty-eight percent of all included cancers were palpable, a situation that does not allow testing the value of early detection. Randomisation was performed at the sites after palpation, while blinding was not guaranteed. In the first round, this “randomisation” assigned 19/24 late stage cancers to the mammography group and only five to the control group, supporting the suspicion of severe errors in the randomisation process. The responsible physicist rated mammography quality as “far below state of

the art of that time”. Radiological advisors resigned during the study due to unacceptable image quality, training, and medical quality assurance. Each described problem may strongly influence the results between study and control groups. Twenty-five years of follow-up cannot heal these fundamental problems. This study is inappropriate for evidence-based conclusions. The technology and quality assurance of the diagnostic chain is shown to be contrary to today's screening programmes, and the results of the CNBSS are not applicable to them.

Key Points

- *The evidence base of the Canadian study (CNBSS) has to be questioned.*
- *Severe flaws in the randomization process and test methods occurred.*
- *Problems were criticized during and after conclusion of the trial by experts.*
- *The results are not applicable to quality-assured screening programs.*
- *The evidence base of this study must be re-analyzed.*

Electronic supplementary material The online version of this article (doi:10.1007/s00330-015-3849-2) contains supplementary material, which is available to authorized users.

✉ S. H. Heywang-Köbrunner
sylvia.heywang@referenzzentrum-muenchen.de

I. Schreer
ingrid.schreer@hamburg.de

A. Hacker
astrid.hacker@referenzzentrum-muenchen.de

M. R. Nofzt
maria.nofzt@uksh.de

A. Katalinic
alexander.katalinic@uksh.de

¹ Ntl. Reference Centre Mammography Munich, Sonnenstr. 29, 80331 Munich, Germany

² Radiologische Allianz Hamburg, Schäferkampsallee 5 – 7, 220357 Hamburg, Germany

³ Institut für Sozialmedizin und Epidemiologie, Universität zu Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany

Keywords Breast cancer · Mammography screening · Canadian National Breast Cancer Screening Study (CNBSS) · Randomization · Image quality

Introduction

The Canadian National Breast Cancer Screening Study (CNBSS) is one of eight large-scale randomized controlled trials on mammography screening. Compared to the other studies, it is an outlier.

Following its first publication in 1992 [1, 2], the quality, randomization, and design of this study was highly debated in the scientific community. The CNBSS is an outlier among the

eight usually cited randomized controlled trials (RCTs), including those with comparable follow-up [3], which showed an average mortality reduction for invited versus non-invited women aged 50–69 years of 23 % and for examined ages about 20 % [4–11].

It contradicts the results of today's quality-assured mammography screening programmes and of the present Canadian screening programme, which report a mortality reduction ranging around 40 % for participating versus nonparticipating women [12–18].

However, it has been classified as one of two studies of high quality in Cochrane reviews, published by Goetzsche since 2000 [6, 19]. A 25-year follow-up of the CNBSS, with added data from two different studies, was published recently [20]. It claimed that mammography screening, performed annually for 5 years (from 1980 to 1985) in the age-group 40–59 years showed no benefit for breast cancer-specific mortality over clinical examination. Taking this outcome seriously, mammography screening should be stopped immediately.

Is this a reasonable conclusion for ongoing mammography screening programs? To answer this question, critical details of the CNBSS, which may have been lost, forgotten, or suppressed, are reconsidered. Another unasked question concerns its applicability to the present situation and to quality-assured screening programmes following European guidelines.

Materials and methods

A systematic search concerning the CNBSS was conducted, which was followed by thematic analysis of study findings and a narrative synthesis. The search included articles from 1981 to 2014 and was performed using PubMed and Embase databases. The following adapted MeSH-Terms were used: “CNBSS”, “Canadian National Breast Screening Study”, “NBSS”, and “Breast cancer”. The selected studies were limited to articles published in English. Papers, comments, editorials, letters, reports, and handbooks dealing with the subject were included.

A total of 240 articles (PubMed n=129, Embase n=111) were identified. Additionally, a manual search and cited reference search generated another 41 articles leading to a total of 281 articles, of which a title and abstract screening was executed by two authors independently. They identified those articles, which treated specific structural, procedural and evaluation issues of the CNBSS. One hundred and twenty-nine potentially relevant articles were saved for full text screening, while 53 articles were duplicates, 97 articles did not meet the inclusion criteria, and two articles were not available. These 129 articles were screened for the following topics: inclusion criteria, randomisation process, contamination of the control group, quality of mammograms, mammographic reading, and diagnostic chain. Forty-three articles were excluded, as they

did not meet these inclusion criteria. Finally, 86 articles were selected for review (Fig. 1).

Results

Sixteen of the 86 eligible articles mentioned ≥ 1 of the debated issues, but argued neither for nor against the quality of the study. Thirty-seven articles of various authors, including 15 articles (co-)authored by five radiological reviewers¹ [21–29] (Kopans, Feig, Logan, Sickles, Tabar) one advisor (Moskowitz), and the responsible physicist [30] (Yaffe) published on “severe flaws” of the study. Thirty-three articles, of which 27 were (co-)authored by the two principal investigators, Miller or Baines, defended the study (Fig. 2). Two of the radiological reviewers resigned during the study.

The following issues were debated (Table 1):

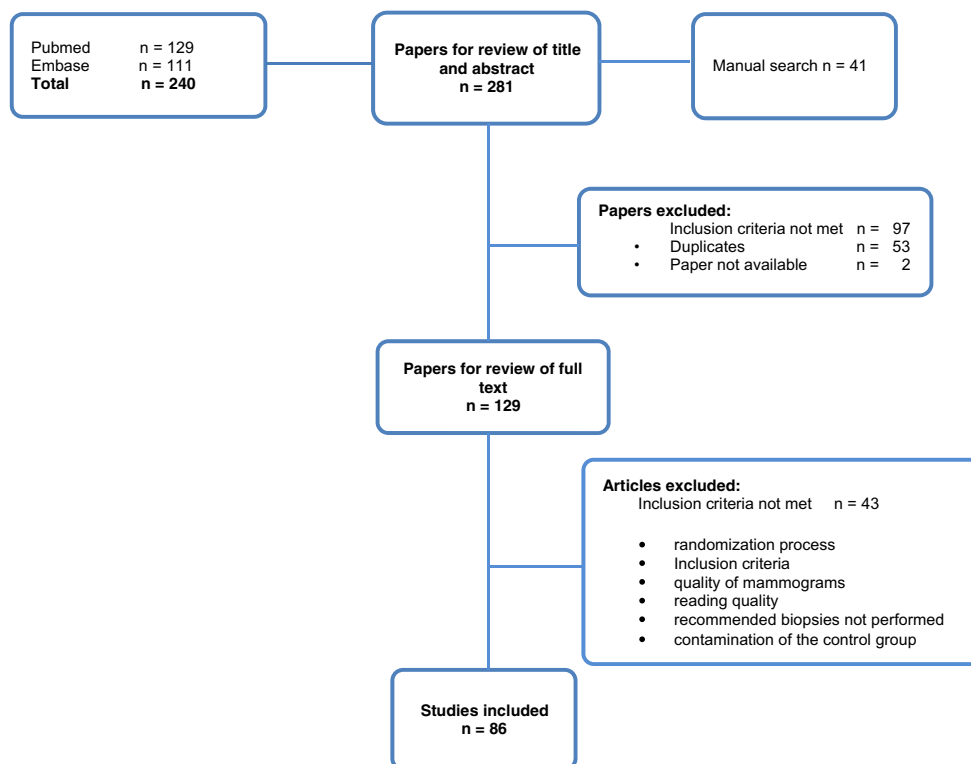
Randomisation: Randomisation was performed *after* a clinical breast examination (CBE) by the principal investigators at each site [28, 31, 32]. This contradicted the initial study design [33], but was documented in the handbook of operation [34] and by an external investigation [32]. While Baines states that “the center coordinators ... were blind to the CBE”, other authors including external reviewers of the study reported the opposite [28, 34]. Also, the coordinator of one of the CNBSS sites was removed from her position because of suspected subversion of the randomisation process [31].

An external investigation [32] for possible fraud stated that in 12 out of 15 sites “...the nurses and probably the coordinators were aware of the findings of the clinical examination when the allocation was made”. In addition, more alterations of the allocation book were found in the mammography screening arm than in the control group (>100 unexplained alterations) [28, 31, 32, 34]. This investigation, which mainly concentrated on checking erasements of participants' names in the allocation book, did “not uncover credible evidence of subversion” [32]. However, Boyd [31] and others (e.g. Kopans, Burhenne) pointed out that various other easy possibilities of subversion existed that could not be excluded with absent blinding.

The external investigation stated, considering that “... referral would not have ensured mammography. The charge has been made that there remained a motive for the examiner or coordinator to subvert the randomisation, if for clinical or other reasons he or she believed that the subject should...have a mammogram” [32].

Unfortunately, a severe imbalance of the distribution of advanced cancers (>4 involved lymph nodes) was noted

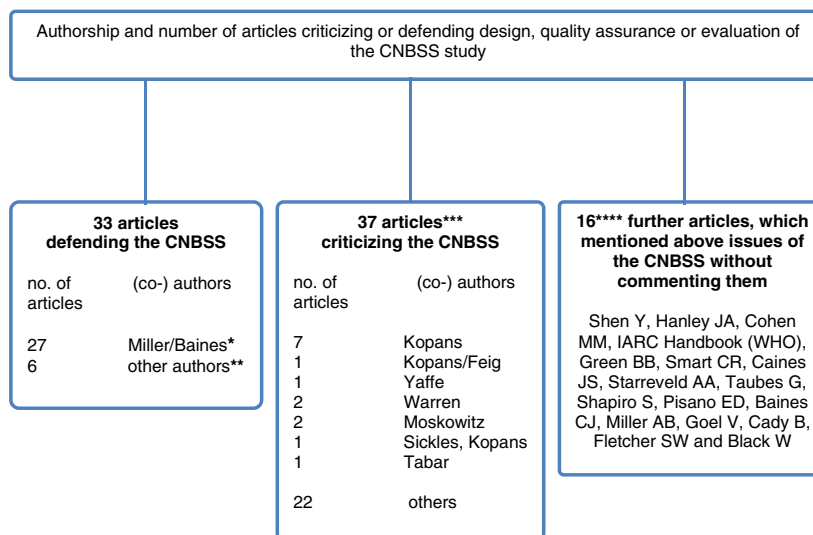
¹ assigned during different time periods

Fig. 1 Flow diagram of systematic literature search

among those women screened in the first round < age 50 years. Nineteen women with far advanced cancers (>4 involved lymph nodes) were allocated to the screening group versus five in the control group, and more women

with prior breast cancer were reported in the screening group (n=8 versus n=1) [28, 31, 34–36].

While Baines argues that more than 50 other variables (“...demographic and risk factors”) showed “virtually

Fig. 2 This chart shows the number of publications which criticized or defended the CNBSS. With very few exceptions, only the main investigators defended the trial. Numerous authors criticized it. Seven of these authors were involved with review or quality assurance of the CNBSS. These authors are mentioned explicitly. Some published several articles concerning the CNBSS. Further publications mentioned some of the issues, but did not comment them. The list of references can be reviewed in the electronic supplementary material (ESM)

* 23 with AB Miller and/or C Baines as first authors,

4 articles with AB Miller and/or C Baines as co-authors

** including articles by Goetzsche [39], Cohen [6], Basinski [60]

*** including 7 persons who were involved in the study, but published severe objections

**** including 1 general article on methodology, important for interpretation of CNBSS data [73]

Table 1 Issues concerning design, quality assurance, and evaluation of the CNBSS: all cited literature

Topics	Arguments	CNBSS	Literature
Randomisation (n=45)	<p>Randomisation was performed at each site after clinical examination (change or violation of initial protocol); on-site randomisation after palpation could no longer guarantee blinding according to independent external review [1]. Various possibilities of subversion existed for anyone working in the study and the probable or possible incentive was to give patients with a highly suspicious finding the best possible diagnosis [2–4].</p> <p>The disproportionally large number of first round participants < age 50 years with advanced cancers entering the mammography group is considered as a strong indicator of flawed randomisation.</p> <p>A low number of late stage cancers, possibly shifted from the control to the study group, could strongly affect and bias the calculation of mortality reduction or overdiagnosis, while other available variables will probably not yet be affected (due to their low association to BC mortality)</p> <p>Defense: “Irrespective of the findings on physical examination...women were independently and blindly assigned randomly” [5]; presence of an incentive is rejected, assuming that symptomatic women would have been seen by a surgeon who “if indicated” [5] could have requested a diagnostic mammogram.</p> <p>They refer to one analysis of a subgroup (Manitoba), which “showed no definitive evidence to support a nonrandom allocation of women with prior breast disease to the mammography arms of the study” [6].</p> <p>They argue that “>50 other variables showed no statistically significant difference between study and control groups” [7].</p> <p>Individual randomisation is assumed generally superior to cluster randomisation</p>	<p>Critique (n=25)</p> <p>Defense (n=14)</p> <p>Not commented, but mentioned (n=7)</p> <p>Critique (n=9)</p> <p>Defense (n=3)</p> <p>Not commented, but mentioned (n=2)</p> <p>Critique (n=28)</p>	<p>[8], [9], [10], [11], [12], [2], [1], [13], [14], [15], [16], [3], [4], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]</p> <p>[29], [26], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [6]</p> <p>[41], [42], [43], [44], 45, [46], [47]</p> <p>[8], [48], [13], [3], [4], [49], [25], [24], [23],</p> <p>[37], [31], [29]</p> <p>[47], [46]</p> <p>[8], [9], [13], [25], [4], [24], [3], [23], [49], [22], [18], [19], [9], [27], [21], [17], [20], [51], [52], [53], [50], [54], [54], [55], [56], [57], [58], [28]</p>
Inclusion criteria (n=14)	<p>By including mainly symptomatic women in this study, the value of screening by definition cannot be tested (68 % of cancers of the mammography arm were palpable).</p> <p>Consequence: dilution of the measurable effect, study is underpowered for testing true screening effect</p> <p>Defense: issue not directly addressed (reasons for choosing this protocol are explained)</p>		
Mammography quality (n=46)	<p>According to the responsible physicist, outdated equipment was used. The equipment at some centers was quite old, and at many centres it lacked key features such as automatic exposure control and grids [30].</p> <p>Image quality was rated by external reviewers to be satisfactory in less than 40 % during 1980-1984 (active recruitment until 1985). According to the external reviewers, improvements of the image quality were regularly demanded and two reviewers resigned during the study due to unacceptable quality. Reported problems concerned incomplete inclusion of the breast tissue, unsharp images, low image contrast, over- or underexposed images resulting in too dark or too bright images, no training of readers, high numbers of obviously missed cancers.</p>		

Table 1 (continued)

Topics	Arguments	CNBSS	Literature
Reading quality (n=26)	The principal investigators state [5], “Facilities and equipment for modern film screen mammography were prerequisites. Quality control procedures were established for radiation physics and mammography interpretation.”	Defense (n=10)	[29], [38], [33], [34], [36], [59], [60], [61], [62], [36]
		Not commented, but mentioned (n=10)	[43], [45], [63], [64], [65], [66], [67], [68], [69], [70]
	Radiologists were not trained in reading mammograms. According to the reviewers, the quality of mammography reading was low and an extraordinarily high number of missed cancers occurred within the 1 year intervals.	Critique (n=16)	[25], [4], [24], [3], [23], [18], [9], [27], [17], [51], [52], [50], [54], [55], [56], [57]
		Defense (n=7)	[33], [34], [36], [38], [7], [71], [62]
	Reading quality and training of readers has not been specifically addressed by proponents of the study	Not commented, but mentioned (n=3)	[68], [63], [67]
Recommended biopsies not performed (n=3)	“25 % of needle localization were recommended but not performed” [4, 5]	Critique (n=1)	[4]
	“...at least one physician refused to do a biopsy on nonpalpable (=mammographically detected) lesions...” [4]		
	“Study surgeons decided if diagnostic follow-up was recommended.” “Most biopsies were done” [36]	Defense (n=2)	[36], [5]
Contamination (n=9)	26 % of women allocated to the control group underwent mammography during the study period. This is a known problem, but may be more pronounced in the trials with individual randomisation. It falsely dilutes the measurable effect	Not commented, but mentioned (n=0)	
		Critique (n=5)	[24], [48], [13], [72], [21]
	Baines argues that a problem concerning 26 % of the control group can only exert a “small” effect	Defense (n=1)	[36]
		Not commented, but mentioned (n=3)	[66], [47], [41]

This table gives an overview of the main issues discussed concerning the CNBSS. The list of references can be reviewed in the electronic supplementary material (ESM)

identical distribution across control and study groups”, Kopans pointed out that “...shifting a much smaller number of advanced cancers to the study group would substantially affect mortality.....without producing a demographic imbalance”.

A bias of randomisation must also be suspected when comparing the hazard ratios of the mammography arm in the prevalence round versus subsequent rounds: 1.47 versus 0.9 [20].

Goetzsche investigated none of the above issues of the randomisation process. Instead, he classified the CNBSS as one of two “high quality” RCTs using one formal criterion: he rated individual randomisation higher than cluster randomisation (where demographic regions are invited/not invited as opposed to individuals).

Comparable to the Malmö study, which also used individual randomisation, a high proportion of cross-over was reported: 26 % of the women in the control group underwent mammography [37].

Inclusion criteria: Women with palpable tumours were included in the CNBSS [1, 2]. This fact alone calls for scrutinizing the inclusion criteria. In the mammography arm of the CNBSS 68 % of breast cancers were already palpable. Furthermore, women with previous breast cancers were included [28, 31], participants whose prognosis may have already be determined by the prior disease.

Quality of the screening procedure: Miller claimed that all equipment was new and that a quality assurance was in place [11, 6]. The responsible physicist

stated [30], “quality was far below state of the art – even at that time (early 1980s). Problems ... from inadequate equipment, ... inappropriate imaging technique and lack of...specialized training....”. This was confirmed by several others including external reviewers [28, 29, 34], who specified that “training for mammography technologists or radiologists and a quality assurance were not present”. The study had initially been overseen by two highly renowned advisors (S. Feig and W. A. Logan). They resigned from the study after 3 years due to unacceptable quality. Subsequent external review rated image quality as acceptable in <40 % of the cases during 1980–1984, in 1985 in about 60 %, and in 1986–7 in <85 %. Problems included incomplete inclusion of the breast tissue, unsharp images, low image contrast, over- or underexposed images [28, 38] (Fig. 3). Overall, a very large number of interval cancers (143/575) resulted in this annual screening trial. It exceeds the rate, which in modern quality-assured mammography screening is considered acceptable for year one of the interval, by a factor of about 3. According to the reviewers, the above problems of poor image quality, inadequate training of personnel and readers may have accounted for the majority of the excess interval cancers. A subgroup analysis by Moskowitz [39] reported an increase of the detection rate with

improving image quality at the end of the study, supporting a correlation with image quality.

Finally, it has been estimated that 25 % of the recommended biopsies were not performed. One of the surgeons was convinced that non palpable lesions detected by mammography did not require biopsy [1, 28].

Discussion and conclusion

The above summary of the published results demonstrates several debatable issues concerning the CNBSS:

- It has to be emphasized that by definition, screening should address asymptomatic women [40]. Current screening is population-based and aims to invite asymptomatic women. The prevalent round of the Canadian trial had a very high proportion of palpable cancers (partly skewed by the nature of their recruitment strategy). Thus, their results are not applicable to current practice where there is uniform access to high quality symptomatic services for women with symptoms. By accepting palpable tumours (most probably advanced and with worse prognosis), the results are skewed: the overall numbers of cancers are artificially high while the palpable cancers cannot contribute to an improved mortality reduction. Thus,

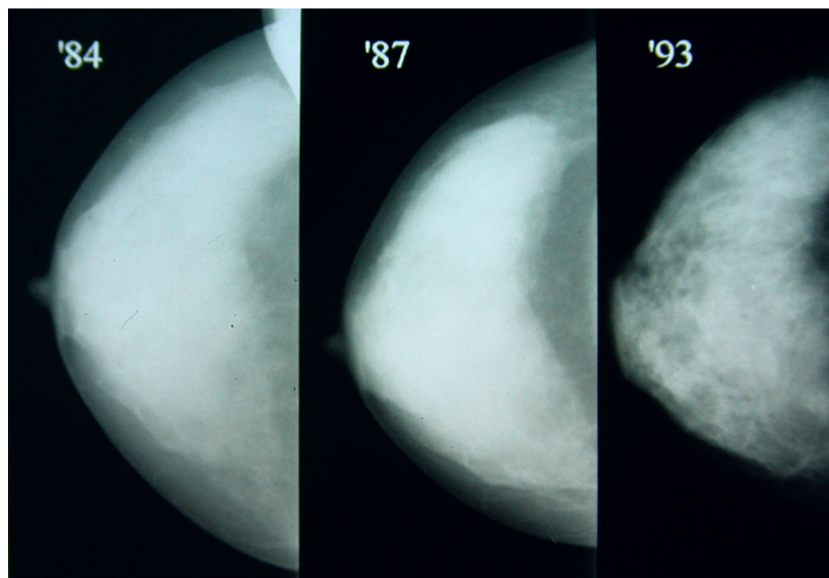


Fig. 3 Images demonstrating the significant changes of technology between 1984, 1987 (CNBSS-study), and a follow-up mammogram of the same patient of 1993. Even though the later technology is still far from present contrast resolution, it becomes obvious that on the former mammograms almost no structures can be discerned

in 80 % of the breast making detection of both masses and microcalcifications almost impossible. Images reproduced courtesy of Dr. Roberta Jong, Sunnybrook Health Sciences Centre, Toronto Canada

the screening effect will be considerably diluted and underestimated. This leads also to underpowered statistics [28, 31].

The documented process of randomisation did not warrant blinding. Thus, any person involved in the study could subvert the randomisation. Also, the probability of subversion was enhanced since mammography was not necessarily offered to women in the control group. We do not assume that the principal investigators committed any fraud. However, they could not have prevented subversion with the chosen protocol. The disproportional distribution of far advanced stages (cancers with >4 involved lymph nodes) in the prevalence screen < age 50 years is highly significant and supports the doubts concerning correct randomisation. An even distribution of demographic and risk factors cannot exclude a bias toward late stage cancers, which may severely impact on the assessment of mortality reduction and calculation of overdiagnoses.

- Long-term mortality reduction was calculated from a maximum of five annual rounds. Because of the short overall duration and continuing entrance of first round screenees, the maximum screening effect could not be reached for many of the participants. Mortality reduction was calculated based on cumulative rates of a mixed trial participation of one to five rounds during up to 5 years. This might lead to a substantial underestimation of the true screening effect compared to a screening programme following approved guidelines (in which participants undergo approximately 10 complete rounds in 20 years) [41].
- The higher evidence classification of individual versus cluster randomized studies is correct in principle. But for screening trials, where non-invited and invited women cannot be blinded, individual randomisation may lead to a much higher contamination of the control group than in a cluster randomized setting. Thus, in the CNBSS, as in other individual randomized screening trials, a substantial underestimation of the screening effect through contamination cannot be ruled out.
- The fact that none of the radiological reviewers including the responsible physicist considered the quality sufficient is highly concerning, as is the described lack of technologist and reader training and the high rate of interval cancers. Two reviewers resigned during the study. How can a method be tested if it is not properly performed and interpreted? What is the value of the results?
- The fact that recommended biopsies of mammographically detected abnormalities were not systematically performed is likely to have distorted the results. What effect is

expected from early detection if suspicious findings are not followed by adequate assessment and therapy?

Both obvious and probable protocol deficiencies are likely to have had an impact on the results counteracting a possible effect of mammography screening on breast cancer mortality and distorting estimates on overdiagnosis. As appropriate randomisation is one of the key validity criteria for RCTs, a study with that kind of violation should not be rated as a high quality study. Whether evidence from such a trial can be used at all must be questioned.

Even if the raised concerns were insignificant and the results were valid, the question remains whether the results and conclusions from a screening trial performed in 1980–85 are applicable to or useful for the assessment of present screening programmes. The most appropriate answer is obviously “no”.

First, the age range in the CNBSS was 40–59 years, which does not apply to the age range of most mammography screening programs today (50–69 or 74 years), as recommended in National and European Guidelines. Because of the lower incidence of breast cancer at younger ages, the absolute effect is lower. Today we know that mammography quality is even more important in the younger age group due to more difficult detection within dense breast tissue. Secondly, the mammographic technique and quality assurance of the complete chain from screening to screen reading, assessment, and treatment in modern population-based mammography screening programs is almost completely different from the CNBSS [42]. Abnormalities are routinely assessed using state of the art minimally invasive methods, and treatment is increasingly standardized and adapted to the stage at detection and the aggressiveness of the cancer. Finally the improved quality today clearly has increased the sensitivity and specificity of mammography screening. The result of 68 % of palpable breast cancers in the mammography arm (average size 1.9 cm) would today be unacceptable for annual (!) screening.

What can we conclude from this short review?

Probable deficits of randomisation and of proper application of the test cannot be repaired by performing a follow-up study. The results of such a study remain biased.

We do not want to discount the CNBSS which represents an enormous and exceptional effort in the 1980s, and we took note of its size, the time, and circumstances when it was conducted. Also, the important question asked in this trial was different from all other screening trials.

However, the chosen methodology led to obvious biases with significant impact on the results, especially when long-term results are considered. Furthermore, it is more than obvious that the setting of mammography in the CNBSS is not comparable to present mammography screening programs.

Therefore, using the CNBSS as “highest evidence” to assess the effects of modern mammography screening programs of the new millennium is not scientifically justified.

Considering the fact that properly performed cohort studies and nested case-control studies with appropriate consideration of length time bias demonstrate a much higher effect of mortality reduction the assumption that the null effect of the CNBSS is “due to availability of chemotherapy” [20] is unproven and highly speculative.

It is an unanswered question, why high-ranking journals, like the BMJ, and representatives of evidence-based medicine close their eyes to these arguments [43] and still refer to the CNBSS as “superior” evidence (against mammography screening).

Taking the CNBSS as an example, the authors want to point out how evidence in the field of breast cancer screening has systematically been omitted, distorted, or inappropriately used over the last decades.

When using CNBSS data, opponents of screening mammography [6, 19, 44, 45] ignore or misinterpret an important part of the existing evidence. The consequence of this recommendation is “waiting until a cancer becomes palpable”. This means that contrary to early detection, women would present at a stage that usually requires aggressive treatment including chemotherapy and more often axillary dissection. There is no doubt that evidence shows that the earlier the stage of breast cancer at diagnosis, the better the prognosis.

In conclusion, the comparison of the settings of the CNBSS with the setting of modern mammography screening is akin to comparing apples with pears. Drawing conclusions from the CNBSS for today’s quality-assured population based screening programmes is an act of negligence.

What we need today is the continuous evaluation of the ongoing mammography screening programmes, including, but not only, breast cancer mortality as an outcome.

Acknowledgements The scientific guarantor of this publication is Sylvia H. Heywang-Koebrunner. The authors of this manuscript declare relationships with the following companies:

Prof. Sylvia Heywang-Koebrunner:

1. I am head of the Reference Center Mammography Screening Munich which is responsible for quality assurance and training of screeners in Bavaria and Thuringia. This makes 50 % of my work.

2. I work as head of a screening unit and I work in private practice, specialized in breast imaging and interventions. About 70 % of my work is associated with mammography screening

Prof. Ingrid Schreer: No other relationships/conditions/circumstances that present a potential conflict of interest

Dr. Astrid Hacker: I'm an employee of the Reference Center for Mammography Screening Munich.

Dr. Maria Nofzt: No other relationships/conditions/circumstances that present a potential conflict of interest

Prof. Dr. Alexander Katalinic: I am chair of the scientific committee of the German Mammography Screening. This is an independent scientific committee, which is giving advice for the ongoing program. The task is honorary; there is no fee paid to me or any committee member. Only costs for travel to the committee meetings are paid by the screening program.

The authors state that this work has not received any funding. One of the authors has significant statistical expertise. Institutional Review Board approval was obtained. No study subjects or cohorts have been previously reported in any other journal. Methodology: The special report is a retrospective review on an existing trial performed in the 1980s (CNBSS) and a lately published follow-up study of the CNBSS. It is performed at one institution.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Miller AB, Baines CJ, To T, Wall C (1992) Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *CMAJ* 147:1459–1476
2. Miller AB, Baines CJ, To T, Wall C (1992) Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *CMAJ* 147:1477–1488
3. Tabár L, Vitak B, Chen TH-H et al (2011) Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 260:658–663
4. World Health Organization, Geneva (2014) WHO position paper on mammography screening. http://apps.who.int/iris/bitstream/10665/137339/1/9789241507936_eng.pdf?ua=1&ua=1. Accessed 2 Jan 2015
5. Duffy SW, Yen AM-F, Chen TH-H et al (2012) Long-term benefits of breast screening. *Breast Cancer Manage* 1:31–38
6. Gotzsche PC, Jørgensen KJ (2013) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 6, CD001877. doi:10.1002/14651858.CD001877.pub5
7. Health Council of the Netherlands (2014) Population screening for breast cancer: expectations and development. http://gr.nl/sites/default/files/201401epopulation_screening_for_breast_cancer_expectations_and_developments_0.pdf. Accessed 5 Nov 2014
8. Independent UK Panel on Breast Cancer Screening (2012) The benefits and harms of breast cancer screening: an independent review. *Lancet* 380:1778–1786
9. Nelson HD, Tyne K, Naik A et al (2009) Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 151:W237–W242
10. Nickson C, Mason KE, English DR, Kavanagh AM (2012) Mammographic screening and breast cancer mortality: a case-control study and meta-analysis. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 21: 1479–1488
11. Pace LE, Keating NL (2014) A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA* 311: 1327–1335
12. Paci E, Working Group EUROSCREEN (2012) Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J Med Screen* 19:5–13
13. Broeders M, Moss S, Nystrom L et al (2012) The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen* 19:14–25

14. Coldman A, Phillips N, Wilson C et al (2014) Pan-Canadian study of mammography screening and mortality from breast cancer. *J Natl Cancer Inst*. doi:10.1093/jnci/dju261
15. Hofvind S, Ursin G, Tretli S et al (2013) Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. *Cancer* 119:3106–3112
16. Olsen AH, Lynge E, Njor SH et al (2013) Breast cancer mortality in Norway after the introduction of mammography screening. *Int J Cancer* 132:208–214
17. Weedon-Fekjær H, Romundstad PR, Vatten LJ (2014) Modern mammography screening and breast cancer mortality: population study. *BMJ* 348:g3701
18. Christiansen P, Vejborg I, Kroman N et al (2014) Position paper: breast cancer screening, diagnosis, and treatment in Denmark. *Acta Oncol (Stockh)* 53:433–444
19. Gotzsche PC, Olsen O (2000) Is screening for breast cancer with mammography justifiable? *Lancet* 355:129–134
20. Miller AB, Wall C, Baines CJ et al (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
21. Kopans DB (1997) Canadian National Breast Screening Study. *Lancet* 350:810
22. Kopans DB (1997) NBSS: opportunity to compromise the process. *CMAJ* 157:247–248
23. Kopans DB (2009) Why the critics of screening mammography are wrong | Diagnostic Imaging. <http://www.diagnosticimaging.com/articles/why-critics-screening-mammography-are-wrong>. Accessed 15 Jan 2015
24. Kopans DB (1993) Mammography. *Lancet* 341:957
25. Sickles EA, Kopans DB (1993) Deficiencies in the analysis of breast cancer screening data. *J Natl Cancer Inst* 85:1621–1624
26. Kopans DB (1990) The Canadian screening program: a different perspective. *Am J Roentgenol* 155:748–749
27. Kopans DB (1992) Response. *J Natl Cancer Inst* 84:1367–1368
28. Kopans DB, Feig SA (1993) The Canadian National Breast Screening Study: a critical review. *AJR Am J Roentgenol* 161:755–760
29. Tabar L, Hsiu-Hsi Chen T (2014) We do not want to go back to the dark ages of breast screening. <http://www.auntminnie.com/index.aspx?sec=ser&sub=def&pag=dis&ItemID=106498>. Accessed 19 Dec 2014
30. Yaffe MJ (1993) Correction: Canada study (Letter to the Editor). *J Natl Cancer Inst* 85(2):94
31. Boyd NF (1997) The review of randomization in the Canadian National Breast Screening Study. Is the debate over? *CMAJ* 156:207–209
32. Bailar JC, MacMahon B (1997) Randomization in the Canadian National Breast Screening Study: a review for evidence of subversion. *CMAJ* 156:193–199
33. Miller AB, Howe GR, Wall C (1981) The National Study of Breast Cancer Screening Protocol for a Canadian Randomized Controlled trial of screening for breast cancer in women. *Clin Invest Med* 4:227–258
34. Burhenne LJ, Burhenne HJ (1993) The Canadian National Breast Screening Study: a Canadian critique. *AJR Am J Roentgenol* 161:761–763
35. Tarone RE (1995) The excess of patients with advanced breast cancer in young women screened with mammography in the Canadian National Breast Screening Study. *Cancer* 75:997–1003
36. Boyd NF, Jong RA, Yaffe MJ et al (1993) A critical appraisal of the Canadian National Breast Cancer Screening Study. *Radiology* 189:661–663
37. Fletcher SW, Black W, Harris R et al (1993) Report of the International Workshop on Screening for Breast Cancer. *J Natl Cancer Inst* 85:1644–1656
38. Baines CJ, Miller AB, Kopans DB et al (1990) Canadian National Breast Screening Study: assessment of technical quality by external review. *AJR Am J Roentgenol* 155:743–747, discussion 748–749
39. Moskowitz M (1992) Response. *J Natl Cancer Inst* 84:1368–1371
40. Wilson JMG, Jungner G (1964) Principles and practice of screening for disease. http://whqlibdoc.who.int/php/who_php_34.pdf. Accessed 3 Nov 2014
41. Hanley JA, McGregor M, Liu Z et al (2013) Measuring the mortality impact of breast cancer screening. *Can J Public Health* 104:e437–e442
42. Perry N, Broeders M, de Wolf C et al (2007) European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. *Ann Oncol* 19:614–622
43. Mokbel K (2011) The Norwegian study suggests that the BMJ is biased against breast cancer screening | The BMJ. <http://www.bmj.com/rapid-response/2011/11/03/norwegian-study-suggests-bmj-biased-against-breast-cancer-screening>. Accessed 19 Dec 2014
44. Biller-Andorno N, Jüni P (2014) Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med* 370:1965–1967
45. Fachgremium Swiss Medical Board (2013) Systematisches mammographie-screening. http://www.medical-board.ch/fileadmin/docs/public/mb/Fachberichte/2013-12-15_Bericht_Mammographie_Final_rev.pdf. Accessed 3 Jan 2015