Advancing research:
One cell at a time
One scientist at a time
One discovery at a time

Benefits and harms of breast cancer screening with mammography in women aged 40–49 years: A systematic review

Caroline van den Ende[1], Anouk M. Oordt-Speets[1], Hilde Vroling[1] and Heleen M. E. van Agt [2]

[1] Pallas, Health Research and Consultancy B.V., Rotterdam, the Netherlands
[2] Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

Early detection of breast cancer through screening can lower breast cancer mortality rates and reduce the burden of this disease in the population. In most western countries, mammography screening starting from age 50 is recommended. However, there is debate about whether breast cancer screening should be extended to younger women. This systematic review provides an overview of the evidence from RCTs on the benefits and harms of breast cancer screening with mammography in women aged 40–49 years. The quality of the evidence for each outcome was appraised using the GRADE approach. Four articles reporting on two different trials—the Age trial and the Canadian National Breast Screening Study-I (CNBSS-I)—were included. The results showed no significant effect on breast cancer mortality (Age trial: RR 0.93 (95% CI 0.80–1.09); CNBSS-I: HR 1.10 (95% CI 0.86–1.40)) nor on all-cause mortality (RR 0.98, 95% CI 0.93–1.03) in women aged 40–49 years offered screening. Among regularly attending women, the cumulative risk of experiencing a false-positive recall was 20.5%. Over-diagnosis of invasive breast cancer at 5 years post-cessation of screening for women aged 40–49 years was estimated to be 32% and at 20 years post-cessation of screening to be 48%. Including ductal carcinoma *in situ*, these numbers were 41% and 55%. Based on the current evidence from randomised trials, extending mammography screening to younger age groups cannot be recommended. However, there were limitations including relatively low sensitivity of screening and screening attendance, insufficient power, and contamination, which may explain the nonsignificant results.

Breast cancer is currently both the most frequent cancer and the most frequent cause of cancer deaths in women in Europe.[1] In 2012, 3.45 million new cases of cancer (excluding non-melanoma skin cancer) and 1.75 million deaths from cancer were estimated in Europe, among which 464,000 new cases of breast cancer and 131,000 deaths from breast cancer.[1] Incidence and mortality rates from breast cancer are expected to rise as a result of the aging population.

Early detection of breast cancer through screening, effective diagnostic pathways, and optimal treatment have the ability to lower current breast cancer mortality rates and reduce the burden of this disease in the population.[2] Many western countries have implemented mammography screening for early detection and treatment of breast cancer to reduce breast cancer mortality. In 2003, the European Parliament promoted the provision of breast cancer screening for all women aged 50–69 years every 2 years.[3] Also in most other western countries, such as the United States (US) and Canada, biennially or triennially mammography screening starting from age 50 is recommended.[4,5] However, there is debate about whether breast cancer screening should be extended to younger women (*i.e.*, 40–49 years).

In effective breast cancer screening programmes, the benefits should outweigh the harms for the population as a whole. Positive effects (benefits) comprise the reduction of breast cancer mortality, reduction of treatments for advanced disease, and reduction of intensive or mutilating treatments. A randomised controlled trial (RCT) is the only method to assess the effects of screening in an unbiased way. Previous meta-analyses of RCTs showed that screening of women aged 39–49 years is associated with a significant reduction in breast cancer mortality of 15–18%.[6–9] However, most of the RCTs included in these analyses were not designed specifically to assess the effect of screening before the age of 50

years, as not all women that participated were younger than 50 years at study entry.

Negative effects (harms) of screening include radiation exposure from mammography, pain during the mammography procedure, consequences of false-positive and false-negative tests, and the occurrence of over-diagnosis. An over-diagnosed cancer, either invasive or non-invasive, is one diagnosed by screening, which would not otherwise have come to attention in the woman's lifetime. As a consequence, women run the risk of decreased quality of life and adverse outcomes of surgery, radiation, and other unnecessary treatments caused by over-diagnosis such as hormonal therapy and chemotherapy. Younger women may benefit less from mammography screening because of factors associated with younger age, including a lower breast cancer incidence and a lower test sensitivity of mammography due to higher breast density and, possibly, faster growing tumours.[10] Some studies estimated the negative effects of screening before the age of 50 years by using modelling techniques.[10,11]

Hence, the positive and negative effects of mammography screening in women aged 40–49 years are still unclear, and there is no consensus on whether or not to offer screening to women in this age group. The latest guideline of the American Cancer Society states that all women should begin having yearly mammograms by age 45, and can change to having mammograms every other year beginning at age 55. Furthermore, women should have the choice to start with yearly mammograms at age 40 if they want to.[12] According to the US Preventive Services Task Force, the decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take into account patient context, including the patient's values regarding specific benefits and harms.[13] Also in Europe, there is no overall agreement. Most European national screening programmes do not invite women younger than age 50, but in some areas in the United Kingdom (UK) women aged 47–49 years now receive invitations for screening as part of a study looking at whether to extend the breast cancer screening age range.[14]

Estimates of the positive and negative effects of mammography screening in women aged 40–49 years based on the latest evidence are required to guide and help policy makers in their decision-making about implementation of the extension of current breast cancer screening programmes. These estimates enable them to make up the balance between the positive effects and negative effects of the extension of breast cancer screening programmes to women aged 40–49 years.

Therefore, the aim of this systematic review is to provide an overview of the evidence on the benefits and harms of breast cancer screening with mammography in women aged 40–49 years. The following research question was investigated: For women aged 40–49 years who are asymptomatic and are not currently under treatment for breast cancer, will screening with mammography as compared to no breast cancer screening with mammography, decrease mortality from breast cancer and what will be the negative effects in terms of false-positive results, false-negative results, chance of over-diagnosis and risk of radiation from mammography?

Material and Methods

A systematic literature review was conducted following a review protocol, the Cochrane guidelines[15] and Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines.[16] The population, intervention, control intervention and the critical outcomes (PICO) were defined by the authors prior to the literature search. The review protocol can be retrieved by contacting the corresponding author.

Search strategy

The literature search was performed in the electronic databases Embase, Medline (OvidSP), Cochrane Library and PubMed from inception to 21 February 2017, by combining search strings for mammography, breast cancer screening and women in the age category of 40–49 years. Only RCTs published in English language were searched. Limits were: no conference abstracts, conference papers, letters or editorials. The complete search strategies are shown in the Appendix.

Eligibility criteria

Studies had to meet the following inclusion criteria: (1) RCTs designed to estimate the benefits/harms of breast cancer screening in women aged 40–49 years from the general population (i.e., the study population at entry includes only women younger than 50 years); (2) intervention/exposure: (any type of) mammography screening (versus no screening); (3) follow-up time of at least 10 years after randomisation; (4) sample size of at least 40,000[17]; (5) disease: primary breast cancer and (6) outcomes: relative reduction in breast cancer-related mortality or all-cause mortality, or proportions of negative effects due to breast cancer screening with mammography (proportion of false-positive/false-negative results, chance of over-diagnosis of breast cancer, risk of radiation). Furthermore, to prevent inclusion of multiple publications on the same study, only the most recent or most complete publication for each dataset for a specific outcome was selected.

Study selection

Articles were selected by screening the titles and abstracts, followed by screening of the full-text articles. The title and abstract selection and subsequent screening of the full-text articles was done in duplicate by two independent researchers. The results were compared and discussed; any doubts or disagreements were resolved by a third researcher. The process of selection and inclusion and exclusion of articles, including the reasons for exclusion of full-text papers, was registered in an Endnote library (version X7.3.1).

Data extraction

Data from included studies were extracted into pre-defined evidence tables by one researcher, in close collaboration with

a second researcher. In case of any doubts or disagreements, a third researcher was consulted. The evidence tables contained information on study characteristics (*i.e.*, country, design, inclusion and follow-up period); study population (*i.e.*, setting, inclusion and exclusion criteria, enrolment age) and study groups (*i.e.*, intervention and comparison, screening protocol, sample size). The outcomes extracted from included studies comprised the reduction of breast cancer mortality and all-cause mortality in the intervention group relative to the control group and the proportions of negative effects due to breast cancer screening with mammography (*e.g.*, the proportion of false-positive/false-negative results). The evidence tables also included a column with comments on risk of bias and other quality aspects of the study.

Quality assessment

The quality of the total body of evidence for each outcome was critically appraised using the GRADE approach (Grading of Recommendations Assessment, Development and Evaluation[1]).

In brief, according to GRADE, evidence on the effects of an intervention can be classified as high, moderate, low and very low. Bodies of evidence from RCTs start as high-quality evidence, whereas those from observational studies start as low-quality evidence. According to a set of predefined criteria, involving within-study risk of bias, directness of evidence, heterogeneity and precision of effect estimates, evidence quality can be increased or decreased. GRADE tables (evidence profiles and summary of findings tables) were created using standard GRADE formats and procedures (with GRADE-pro[18]) to summarise these quality aspects and other specific details of the included studies, such as study outcomes. The tables were compiled by one researcher in close collaboration with a second researcher, and if necessary a third researcher.

Synthesis of results

Pooling of data was planned if more than one study on a given outcome was available and data from these studies were sufficiently homogeneous in terms of clinical, methodological and statistical characteristics. Otherwise, narrative syntheses were conducted.

Results

The search yielded 2,042 unique records in the electronic databases, of which the full text of 70 articles was assessed and four articles were finally included after applying the inclusion and exclusion criteria. A flow diagram of the selection process is presented in Figure 1.

Characteristics of included studies

Two of the included articles presented data from the UK Age trial,[19,20] which was undertaken in 23 breast-screening units in England, Wales and Scotland. From 1991 to 1997, women of age 39–41 years were included and randomised to receive mammography or usual care. Screening in the trial was by two-view mammography at first screen and by single view thereafter. The women were identified from lists of patients of general (family) practitioners held on local Health Authority databases. Follow-up lasted to December 31, 2011. One of the two studies assessed the reduction on breast cancer mortality and all-cause mortality,[20] whereas the other reported on the frequency of false-positive screens.[19]

The other two included articles described results from the Canadian National Breast Screening Study-I (CNBSS-I), which involved women aged 40–49 years who were recruited between January 1980 and March 1985 through a general publicity campaign and review of population lists.[21,22] In this study, all women received initial breast physical examination and instruction on breast self-examination before randomisation to two-view mammography or usual care. In one of the studies, the women were followed until age 60 for mortality from breast cancer (all women had reached age 60 before completion of follow-up in December 2005).[22] The other study reported estimates of over-diagnosis using post-screening cessation cut-off points from 1 to 20 years.[21] The characteristics of the included studies, and the reported outcomes and risk of bias in the studies, are summarised in detail in Table 1.

Reduction in breast cancer mortality

In the study of Moss *et al.* (Age trial), the reduction in breast cancer mortality in the screening group relative to the control group was 7% at a median follow-up of 17.7 years, which did not reach statistical significance (RR 0.93 (95% CI 0.80–1.09)).[20] The hazard ratio (HR) for death from breast cancer before the age of 60 years, given the use of screening mammography, was 1.10 (95% CI 0.86–1.40; $p = 0.45$) in the study of Narod *et al.* (CNBSS-I).[22] In both studies, the 95% CI around the effect estimate included both no effect and appreciable benefit. Accordingly, the quality of evidence was downgraded to moderate due to imprecision (see Table 2 for the complete GRADE evidence profile). Owing to the heterogeneity of the data (*i.e.*, different outcome measures), it was not possible to provide summarising outcome measures or to conduct a meta-analysis.

Reduction in all-cause mortality

The reduction in all-cause mortality in the screening arm relative to the control arm was 2% in the study of Moss *et al.* (Age trial), however, not statistically significant (RR 0.98, 95% CI 0.93–1.03).[20] As the trial was not powered to detect an effect on all-cause mortality, the quality of evidence was downgraded to moderate due to risk of bias (Table 2).

Frequency of false positives

Of the 53,884 women randomised to the intervention group in the study of Johns *et al.* (Age trial), 7,893 women (14.6%) experienced one or more false-positive screens (defined as routine trial screens where initial mammographic findings led to recall for additional procedures, but further assessment did
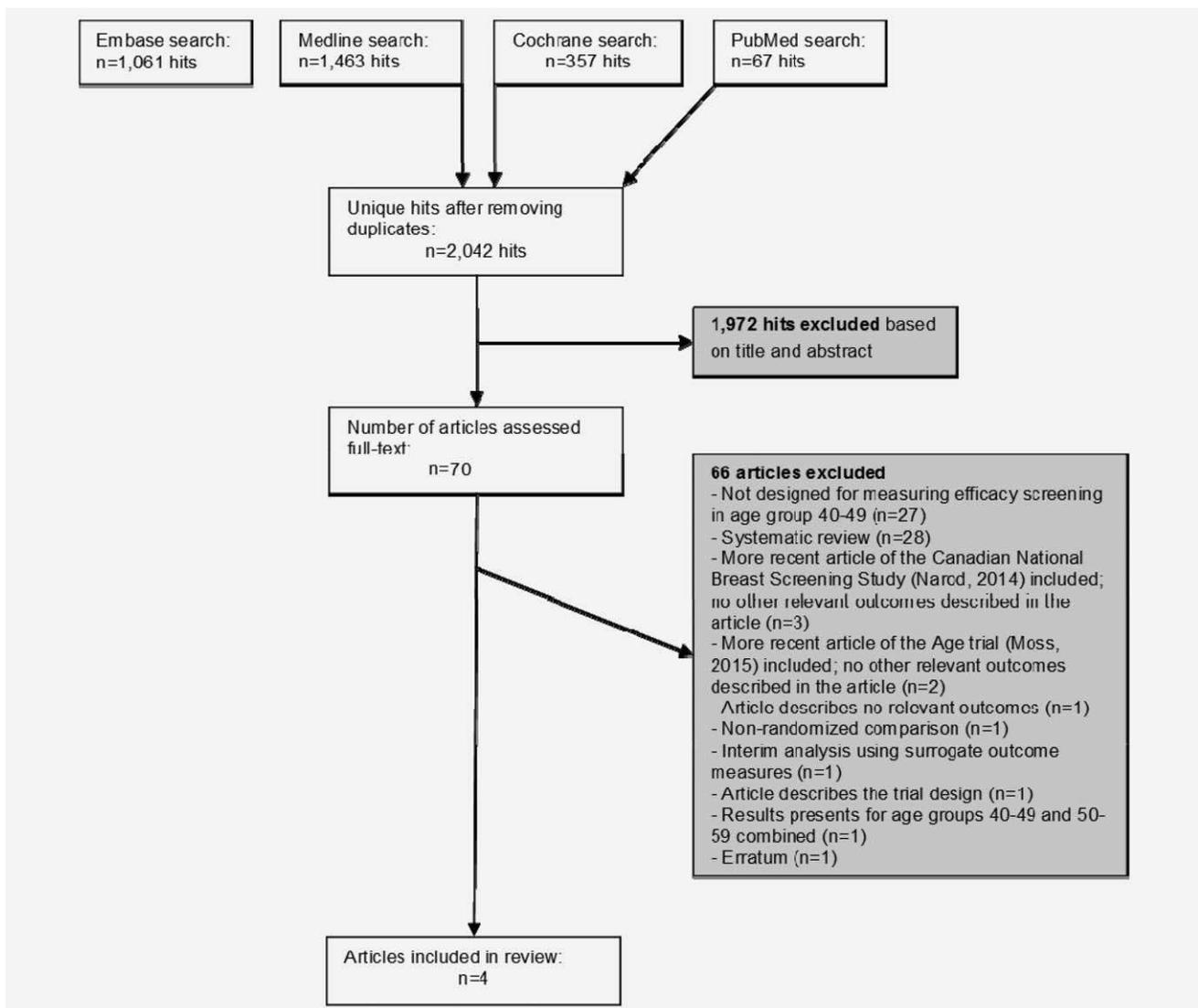
Figure 1. Flowchart showing the selection process of the systematic review.

not result in a diagnosis of breast cancer at that episode) during the course of the trial.[19] The observed cumulative risk of experiencing a false-positive recall over the first seven screens was 20.5% among regular attenders (those attending ≥7 routine screens during the trial; $n = 23,245$). No downgrading was done for this outcome and therefore the quality of evidence was rated as high (Table 2).

Over-diagnosis

Over-diagnosis of invasive breast cancer (calculated as cumulative breast cancers in the mammography arm after certain years of follow-up minus the cumulative breast cancers in the control arm after certain years of follow-up, divided by the numbers of screen-detected breast cancers during the trial period in the mammography arm) at 5 years post-cessation of screening was estimated to be 32% in the study of Baines et al. (Canadian trial); 30 years post-cessation, this was 48%. If women with ductal carcinoma in situ were also included,

these estimates were 41% and 55%, respectively. The authors estimated that overall, approximately 30% of invasive screen-detected breast cancers in women aged 40–49 years were over-diagnosed.[21]

Discussion

The aim of this systematic review was to compile the evidence on the benefits and harms of breast cancer screening with mammography in women aged 40–49 years. Four articles reporting on two different trials were included. The results showed no statistically significant effect on breast cancer mortality nor on all-cause mortality in women aged 40–49 years offered mammography screening. Whereas one of the two included trials (CNBSS-I) reported a small and non-significant excess of breast cancer deaths before the age of 60 years, the other trial (UK Age trial) showed a non-significant reduction of breast cancer mortality.[20,22] Also the reduction in all-cause mortality found in the Age trial was small and

Table 1. Study characteristics, outcomes and risk of bias

Study, country	Setting, study population and study period	Study groups and sample size	Outcomes	Risk of bias				
				Random sequence generation	Allocation concealment	Blinding	Loss to follow-up/ IT principle observed	Other limitations
Moss, 2015 Age trial UK	23 NHS breast-screening units in England, Wales and Scotland Women 39–41 years *Inclusion period* 1991–1997 *Follow-up* To Dec 31, 2011 (median 17.7 years; IQR 16.8–18.8)	Mammography vs usual care I: $n = 53,883$ C: $n = 106,953$	- Reduction in breast cancer mortality - Reduction in all-cause mortality	- Adequate - Individual randomisation by computer, stratified by GP practice - Randomisation ratio 1:2	- Adequate - Randomisation and allocation was carried out by computer	Not possible for patients and personnel. Information about cause of death was obtained from the NHSCR	Analysis based on intention-to-treat principle I: $n = 650$ lost to follow-up because of emigration C: $n = 1,183$ lost to follow-up because of emigration	The trial was not powered to detect an effect on all-cause mortality
Johns, 2010 Age trial UK	23 NHS breast-screening units in England, Wales and Scotland Women 39–41 years *Inclusion period* 1991–1997 *Follow-up* To Dec 31, 2004 (mean NR)	Mammography vs usual care; this analysis included the intervention arm only $n = 53,884$	Frequency of false-positives*	- Adequate - Individual randomisation by computer, stratified by GP practice - Randomisation ratio 1:2	- Adequate - Randomisation and allocation was carried out by computer	Not possible for patients and personnel	Loss to follow-up NR; analysis based on intention-to-treat principle	None
Baines, 2016 Canadian National Breast Screening Study-I Canada	15 screening centres in 6 Canadian provinces Women 40–49 years *Inclusion period* January 1980–March 1985 *Follow-up* To 25 years post-entry to the study	Mammography and physical breast examination vs usual care I: $n = 25,216$ C: $n = 25,220$	Over-diagnosis	- Adequate - Individual randomisation, stratified by centre and 5-year age group	- Adequate - Use of lists with pre-printed identification numbers and group designations	Not possible for patients and personnel	Loss to follow-up NR; analysis based on intention-to-treat principle	- All women received initial breast physical examination and instruction on breast self-examination before randomisation; therefore the study sample consisted of pre-screened volunteers instead of unselected women

Mini Review

Mini Review

Table 1. Study characteristics, outcomes and risk of bias (Continued)

Study, country	Setting, study population and study period	Study groups and sample size	Outcomes	Risk of bias				Other limitations
				Random sequence generation	Allocation concealment	Blinding	Loss to follow-up/ IT principle observed	
Narod, 2014 Canadian National Breast Screening Study-I Canada	15 screening centres in 6 Canadian provinces Women 40–49 years *Inclusion period* January 1980–March 1985 *Follow-up* To Dec 31, 2005; all women were followed-up to age 60 (mean 15.2 years)	Mammography and physical breast examination vs usual care I: $n = 25,216$ C: $n = 25,220$	Reduction in breast cancer mortality	- Adequate - Individual randomisation, stratified by centre and 5-year age group	- Adequate - Use of lists with pre-printed identification numbers and group designations	Not possible for patients and personnel. Cause of death was assessed blinded	Loss to follow-up NR; analysis based on intention-to-treat principle	All women received initial breast physical examination and instruction on breast self-examination before randomisation; therefore, the study sample consisted of pre-screened volunteers instead of unselected women

Abbreviations: I, intervention; C, control; IT, intention-to-treat; NHS, National Health Service; NHSCR, National Health Service central register; NR, not reported.
False-positive screens were defined as routine trial screens where initial mammographic findings led to recall for additional procedures, but further assessment did not result in a diagnosis of breast cancer at that episode.

did not reach statistical significance.[20] Among regularly attending women, the cumulative risk of experiencing a false-positive recall was 20.5% in the Age trial.[19] Over-diagnosis of invasive breast cancer at 5 years post-completion of CNBSS screening for women aged 40–49 was estimated to be 32%; 20 years post-cessation of screening this was 48%. If women with ductal carcinoma *in situ* were also included, these estimates were higher: 41% and 55%, respectively.[21]

Until now, only few studies have been published that were specifically designed to assess the effect of breast cancer screening in women younger than 50 years. Most mammography screening trials include women of a broader age range (*e.g.*, 40–64 years) and subsequently analyse the results for the smaller sample of women aged 40–49 years. Interpreting stratified results should be done with caution, however. Although the CNBSS-I included women aged 40–49 years, some women reached the age of 50 years shortly after inclusion. It is therefore possible that any benefit or harm results from breast cancer screening taking place after age 50,[23] when women will receive the first invitation in the national screening programme. The Age trial overcomes this issue by inviting women aged 39–41 years at study entry and is thus up to now the only trial that is set up specifically to evaluate the effectiveness of screening in women under 50 years of age.

The authors of the Canadian trial assumed that none of the women assigned to the control arm underwent mammography before age 50; however, it is possible that some of those women underwent mammography off-study before the age of 50, and therefore, cross-over might have reduced the effect of mammography on the true mortality difference. In addition, women in the intervention group underwent on average only four or five screens during the 10-year period of the annual screening program, which may have also contributed to a reduced effect of the screening. On the other hand, the quality of mammography seems adequate, consisting of two view mammography, whereas in the Age trial only, the first screening was a two view mammogram. However, both the method of randomisation and the quality of the mammograms in the Canadian Trial have been questioned.[12,24] Randomisation was done on the basis of lists supplied by the central office with pre-printed identification numbers and group designations[25] and took place after a clinical breast examination (CBE). Therefore there would be some knowledge at the screening site of palpable abnormalities before the official registration of the participant occurred.[12] Although it was stated that the centre coordinators were blinded for CBE,[25] others reported the existence of possibilities to subvert the randomization.[24] Unfortunately, there was an imbalance in the distribution of observed cancers between the screening and control group in the first screening round: respectively 19 versus 5 advanced cancers (>4 involved nymph nodes),[24] which could be the result of not following the randomisation procedure or just chance. In response to this critics, the authors of the Canadian trial re-evaluated the effect of

Table 2. GRADE evidence profile for the benefits and harms of mammography screening in women aged 40–49 years

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Routine mammography	Usual care	Relative (95% CI)	Absolute	Quality	Importance
		Quality assessment					No of patients		Effect			
Reduction in breast cancer mortality												
1[9]	Randomised trial[7]	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious[1]	None	242/53,883 (0.45%)	412/ 106,953 (0.48%)	RR 0.93[2] (0.80–1.09)	337 fewer per 1,000,000 (from 963 fewer to 433 more)	⊕⊕⊕O MODERATE	CRITICAL
1[10]	Randomised trial[8]	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious[3]	None	0.53%[4]	0.48%[4]	HR 1.10 (0.86–1.40), $p = 0.45$	-	⊕⊕⊕O MODERATE	CRITICAL
Reduction in all-cause mortality (median follow-up 17.7 years)												
1[9]	Randomised trial[7]	Serious[5]	No serious inconsistency	No serious indirectness	No serious imprecision	None	2,127/53,883 (3.95%)	4,320/ 106,953 (4.04%)	RR 0.98[6] (0.93–1.03)	808 fewer per 1,000,000 (from 2,827 fewer to 1,212 more)	⊕⊕⊕O MODERATE	CRITICAL
Frequency of false-positives (mean/median/range follow-up not reported)												
1[11]	Randomised trial	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	7,893/53,884 (14.6%)	-	-	-	⊕⊕⊕⊕ HIGH	CRITICAL
Over-diagnosis (during screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	284	225	-	Over-diagnosis 28%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (1 year post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	327	262	-	Over-diagnosis 31%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (2 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	379	308	-	Over-diagnosis 33%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (3 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	435	363	-	Over-diagnosis 34%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (4 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	487	421	-	Over-diagnosis 31%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (5 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	544	476	-	Over-diagnosis 32%[14]	⊕⊕⊕O MODERATE	CRITICAL

Mini Review

Table 2. GRADE evidence profile for the benefits and harms of mammography screening in women aged 40–49 years (Continued)

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Routine mammography	Usual care	Relative (95% CI)	Absolute	Quality	Importance
				Quality assessment			No of patients		Effect			
Over-diagnosis (10 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	912	817	-	Over-diagnosis 45%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (15 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	1,386	1,311	-	Over-diagnosis 35%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (20 years post-screening), invasive cancers only												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	1,725	1,622	-	Over-diagnosis 48%[14]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (during screening), invasive and *in situ* cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	326	234	-	Over-diagnosis 37%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (1 year post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	371	271	-	Over-diagnosis 40%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (2 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	424	318	-	Over-diagnosis 43%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (3 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	480	373	-	Over-diagnosis 43%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (4 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	533	432	-	Over-diagnosis 41%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (5 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	590	487	-	Over-diagnosis 41%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (10 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	958	828	-	Over-diagnosis 52%[15]	⊕⊕⊕O MODERATE	CRITICAL
Over-diagnosis (15 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	1,432	1,322	-	Over-diagnosis 44%[15]	⊕⊕⊕O MODERATE	CRITICAL

Table 2. GRADE evidence profile for the benefits and harms of mammography screening in women aged 40–49 years (Continued)

	Quality assessment						No of patients		Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Routine mammography	Usual care	Relative (95% CI)	Absolute	Quality	Importance
Over-diagnosis (20 years post-screening), invasive and in situ cancers												
1[12]	Randomised trial	Serious[13]	No serious inconsistency	No serious indirectness	No serious imprecision	None	1,771	1,633	-	Over-diagnosis 55%[15]	⊕⊕⊕○ MODERATE	CRITICAL

[1]Wide 95% CI indicates imprecision (the 95% CI around the RR for breast cancer mortality includes both no effect and appreciable benefit).
[2]RR based on rate of breast cancer deaths per 1,000 women years (rate mammography 0.257 per 1,000 women years; rate control 0.276 per 1,000 women years); however, the RR based on persons as denominator is the same (0.93).
[3]Wide 95% CI indicates imprecision (the 95% CI around the HR for breast cancer mortality includes both no effect and appreciable harm).
[4]Absolute numbers of cumulative incidence of death from breast cancer calculation not reported in article (sample size mammography n = 25,216; sample size control, n = 25,220).
[5]The trial was not powered to detect an effect on all-cause mortality.
[6]RR based on rate of all cause deaths per 1,000 women years (rate mammography 2.26 per 1,000 women years; rate control 2.31 per 1,000 women years); however, the RR based on persons as denominator is the same (0.98).
[7]Follow-up median 17.7 years.
[8]Follow-up mean 15.2 years.
[9]Moss, Lancet, 2015.
[10]Narod, Curr Oncology, 2014.
[11]Johns, Cancer Epidemiol Biomarkers Prev, 2010.
[12]Baines, Prevent Med, 2016.
[13]All women received initial breast physical examination and instruction on breast self-examination before randomisation; therefore, the study sample consisted of pre-screened volunteers instead of unselected women.
[14]The numerator is the difference in number of cancers in the mammography arm less than those in the control arm, and the denominator is the 213 screen-detected cancers in the mammography arm.
[15]The numerator is the difference in cancers in the mammography arm compared to the control arm, and the denominator is the 249 screen-detected cancers in the mammography arm.

mammography screening by excluding all prevalent cancers at study entry and found a statistically non-significant 10% reduction in breast cancer mortality (Hazard Ratio in screening rounds 2–5: 0.90 (95% CI 0.69–1.16)).[26] This would be consistent with the statistically non-significant 17% reduction in the Age Trial.

Although the authors of the Canadian Trial reported that mammography was in accordance with standard practice, that facilities and equipment for modern film screen mammography were prerequisites, that quality control procedures were established for radiation physics and mammography interpretation, and that breast examiners received a month of training,[27] critics judged the equipment and mammography images to be below the standard of time and reported that specific training to perform and interpret mammography was lacking.[12,24] However, the authors of the Canadian Trial argue that sensitivity of the mammography employed in the screening centres was representative of the quality of the technology delivered at cancer centres and teaching hospitals and that the screening examination was properly conducted.[27] Previous results of the Age Trial, after a mean follow-up of 10.7 years, showed a non-significant risk reduction for breast cancer mortality (RR 0.83, 95% CI 0.66–1.04).[28] According to the authors, the absolute effect of mammography screening is difficult to assess when deaths from cancers diagnosed after the intervention phase of the trial are included. Analyses restricted to tumours diagnosed in the intervention phase showed a significant reduction of breast cancer mortality in the first 10 years after diagnosis, but there was no significant reduction after 10 years, at the time when both groups received the same care. The authors argue that with increasing time, the effect will be diluted by breast cancers diagnosed after the end of screening, including those detected by screening after age 50 in the national program. Some other issues might also explain the lack of effect of breast cancer screening in their study. One was that the power of the study was diminished due to a smaller sample than planned. This was a result of financial and workload constraints and lower breast cancer mortality in the control group than anticipated, probably due to improvement of treatment and survival since the initial power calculations. Second, women who moved out of the region were not invited any more, and as a result, <55% of women in the intervention group was actually screened by the seventh screening round. The authors argue that the screening effect at later follow-up would be greater in a national screening programme where these women would still have been invited.[20] Modelling studies can be helpful to estimate and understand the effects of varying the total number of women included, follow-up time or proportion of screened women.

The Canadian Trial assessed breast cancer mortality but not all-cause mortality. Narod *et al.* explained that a specific aim of the study was to test the hypothesis that treatment of otherwise indolent breast cancers has the potential to reactivate dormant metastases and to accelerate their growth,[29]

Mini Review

which may lead to a transient increase in the number of deaths from breast cancer.[22] Only the Age trial included all-cause mortality as an outcome measure. According to Gotzsche et al.,[30] assignment of breast cancer mortality is unreliable and biased in favour of screening. However, very large trials are needed to assess the effect of screening on all-cause mortality.

Harms of mammography screening among women aged 40–49 years were evaluated in the Age trial by assessment of false-positive rates. False-positive rates differ widely between countries, mainly depending on the tendency to interpret mammograms as abnormal.[31] Comparisons of false-positive rates between age groups should therefore preferably be made within the country at stake. In the Age trial, false-positive rates at the first and subsequent screenings were 4.9% and 3.2%, which resemble those (7.9% initial rate and 3.2% subsequent rates) in the UK mammography screening programme.[32] Given the different cancer incidence rates between younger and older women, the Positive Predictive Value (PPV) may be a better estimate to compare false-positive results between age groups.[19] The PPV in the Age trial was 2% at first screens and 3–5% in subsequent screens, whereas in the UK screening program, the values of the PPV were higher, 8% and 16%, respectively.[32] In the Age trial, the cumulative risk of a false-positive recall of regularly attending women over seven screens was 20.5% and 28% over 10 screens, which was higher than those 10–11% found in five units in the national screening program after four screening rounds.[33] The higher rate found in the Age trial as compared to screening from 49+ could be explained by the fact that the sensitivity and PPV of mammography decrease with increasing breast density. As younger women have denser breast tissue, which diminishes gradually with increasing age, false-positive results are expected to be higher.[19,31]

Another possible harmful effect of screening women aged 40–49 years concerns an increased amount of over-diagnosis. To estimate over-diagnosis correctly, sufficient follow-up time is needed to allow time for the compensatory drop after the end of the intervention phase.[20] According to the Independent UK Panel on Breast Cancer Screening, there is no single optimum way to calculate over-diagnosis, although they describe the two most useful estimates. These are (a) from the population perspective, the proportion of all cancers ever diagnosed in women invited to screening that are over-diagnosed and (b) from the perspective of a woman invited to screening, the probability that a cancer diagnosed during the screening period represents over-diagnosis.[34] Another method to calculate excess cancers, namely, as a proportion of cancers detected by screening in women invited for screening, was also described by the Panel and was used in one of the publications of the Canadian trial included in this review. In this study, over-diagnosis was estimated up to 25 years post-entry to the trial. The authors state that these estimates may have been confounded by subsequent screening in the population by national breast screening programs, especially

after 10 years as they include over-diagnosis from post-CNBSS screening.[21] In the included publication of the Age trial, the authors conclude that their results provide no evidence that screening in the trial resulted in any over-diagnosis in addition to any occurring as a result of screening in the national program. The long-term incidence of all breast cancers, including those diagnosed after entry to the national program, was somewhat lower in the intervention group.[20]

We did not find any studies on other negative effects of screening, such as the risk of radiation induced breast cancer. According to the authors of the Age trial, the proportion of women for whom the risk might outweigh the benefit of screening is very small. An earlier study of Law and Faulkner estimated that the cancer detection/induction ratio, an index of the benefit-to-risk ratio, would exceed 1. However, in the Age trial, it was estimated that the number of cancers induced per 1,000 women aged 40–49 years screened would be reduced by a factor of around 0.75 (assuming that 5% of screens other than the first are by two views), whereas the detection rates would be around 30% higher. This would lead to an increase in the cancer detection/induction ratio by a factor of 1.7. Nelson et al. also stated that the absolute level of radiation exposure and corresponding radiation risk from mammography is very low.[8]

Other research on breast cancer screening comprises mainly population-based observational studies[35] or secondary analyses of population screening trials. Two meta-analyses (Nelson et al. 2009[8] and Magnus et al. 2011[7]) including screening trials in all age groups, and the Age trial and Canada trial found a breast cancer mortality reduction of 15%[8] and 17%,[7] respectively, in women below 50 years of age. In agreement with another meta-analysis on screening in all age groups,[30] Nelson et al. did not include three of the trials that were included in the Magnus meta-analysis, because of suboptimal randomisation. However, the US Preventive Services Task Force (USPSTF) does not explicitly recommend screening for women younger than 50. They stated "Women aged 40 to 49 years experience the highest rate of additional imaging, whereas their biopsy rate is lower than that for older women. Mammography screening at any age is a trade-off of a continuum of benefits and harms. The ages at which this trade-off becomes acceptable to individuals and society are not clearly resolved by the available evidence".

The included trials in this systematic review concern film mammography. Nowadays, new technologies for mammography have emerged, such as digital mammography which has shown to be more sensitive in younger women with dense breasts as compared to film mammography.[36] A recent modelling study indicated that digital mammography screening in women aged 40–49 years in addition to current screening in the Netherlands is cost-effective.[10] The model predicted 26% reduction of breast cancer mortality with current screening in women aged 50–74 years, which was in line with earlier research. Lowering the starting age of screening

to 40 years would reduce mortality with an additional 5%. An increase in the number of false-positives by 74 per 11,000 (60%) and an increase of over-diagnosis by 0.33 per 1,000 (11%) is also expected. Cost-effectiveness of digital mammography in women aged 40–49 years was shown in a US modelling study as well.[11]

Up to now, the effectiveness of breast cancer screening with mammography for women aged 40–49 years has not been proven in randomised trials. However, there were limitations regarding the power of the study, follow-up time and screening attendance that may explain the non-significant effects. Therefore, based on the current evidence from randomised trials, extending mammography screening to younger age groups cannot be recommended yet. However, modelling studies indicate cost-effectiveness of breast cancer screening with digital mammography in women aged 40–49 years. Further research should focus on trials with new mammography technology.

Acknowledgements

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013; 49: 1374–403. doi: 10.016/j.ejca.2012.12.027. Epub 3 Feb 26.

2. Perry N, Broeders M, de Wolf C, et al. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition–summary document. *Ann Oncol* 2008; 19: 614–22. Epub 2007 Nov 17.

3. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013; CD001877.

4. Baum M. Reflections on screening mammography and the early detection of breast cancer. *Curr Oncol* 2014; 21: 215–6.

5. Loberg M, Lousdal ML, Bretthauer M, et al. Benefits and harms of mammography screening. *Breast Cancer Res* 2015; 17: 63.

6. Hendrick RE, Smith RA, Rutledge 3rd JH, et al. Benefit of screening mammography in women aged 40–49: a new meta-analysis of randomized controlled trials. *J Natl Cancer Inst Monogr* 1997; 87–92.

7. Magnus MC, Ping M, Shen MM, et al. Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39–49 years: a meta-analysis. *J Women Health* 2011; 20: 845–52.

8. Nelson HD, Tyne K, Naik A, et al. Screening for breast cancer: Systematic evidence review update for the US Preventive Services Task Force 2009.

9. Smith RA, Duffy SW, Gabe R, et al. The randomized trials of breast cancer screening: what have we learned? *Radiol Clin North Am* 2004;42: 793–806.

10. Sankatsing VD, Heijnsdijk EA, van Luijt PA, et al. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. *Int J Cancer* 2015; 137: 1990–9. doi: 10.002/ijc.29572. Epub 2015 May 8.

11. Stout NK, Lee SJ, Schechter CB, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. *J Natl Cancer Inst* 2014; 106: dju092.

12. Yaffe MJ. Point: mammography screening-sticking to the science. *Curr Oncol* 2015; 22: 174–6.

13. Gadducci A, Sergiampietri C, Tana R. Alternatives to risk-reducing surgery for ovarian cancer. *Ann Oncol* 2013; 24 (Suppl 8): viii47–53.

14. Bodelon C, Pfeiffer RM, Buys SS, et al. Analysis of serial ovarian volume measurements and incidence of ovarian cancer: implications for pathogenesis. *J Natl Cancer Inst* 2014; 106.

15. Higgins J.P.T. GSe. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].*ed. The Cochrane Collaboration: Available from http://handbook.cochrane.org, 2011.

16. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6: e1000097. doi: 10.1371/journal.pmed. Epub 2009 Jul 21.

17. Moss S, Draper GJ, Hardcastle JD, et al. Calculation of sample size in trials of screening for early diagnosis of disease. *Int J Epidemiol* 1987; 16: 104–10.

18. King-Spohn K, Pilarski R. Beyond BRCA1 and BRCA2. *Curr Probl Cancer* 2014; 38: 235–48.

19. Johns LE, Moss SM, Age Trial Management G. False-positive results in the randomized controlled trial of mammographic screening from age 40 ("Age" trial). *Cancer Epidemiol Biomarkers Prev* 2010; 19: 2758–64.

20. Moss SM, Wale C, Smith R, et al. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. *Lancet Oncol* 2015; 16: 1123–32. doi: 10.016/S470-2045(15)00128-X. Epub 2015 Jul 20.

21. Baines CJ, To T, Miller AB. Revised estimates of overdiagnosis from the Canadian National Breast Screening Study. *Prevent Med* 2016; 90: 66–71.

22. Narod SA, Wall C, Baines C, et al. Impact of screening mammography on mortality from breast cancer before age 60 in women 40 to 49 years of age. *Curr Oncol* 2014; 21: 217–21.

23. de Koning HJ, Boer R, Warmerdam PG, et al. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Natl Cancer Inst* 1995; 87: 1217–23.

24. Heywang-Kobrunner SH, Schreer I, Hacker A, et al. Conclusions for mammography screening after 25-year follow-up of the Canadian National Breast Cancer Screening Study (CNBSS). *Eur Radiol* 2016; 26: 342–50.

25. Miller AB, Baines CJ, To T, et al. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *CMAJ* 1992; 147: 1459–76.

26. Narod SA. Reflections on screening mammography and the early detection of breast cancer: a countercurrents series. *Curr Oncol* 2014; 21: 210–14.

27. 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. *BMJ* 2014; 348: g366.

28. Moss SM, Cuckle H, Evans A, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006; 368: 2053–60.

29. Retsky M, Demicheli R, Hrushesky W, et al. Surgery triggers outgrowth of latent distant disease in breast cancer: an inconvenient truth?. *Cancers (Basel)* 2010; 2: 305–37.

30. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011; 1: CD001877.

31. Otten JD, Fracheboud J, den Heeten GJ, et al. Likelihood of early detection of breast cancer in relation to false-positive risk in life-time mammographic screening: population-based cohort study. *Ann Oncol* 2013; 24: 2501–6. doi: 10.1093/annonc/mdt227. Epub 2013 Jun 19.

32. Bennett RL, Blanks RG, Patnick J, et al. Results from the UK NHS Breast Screening Programme 2000–05. *J Med Screen* 2007; 14: 200–4.

33. Wallis M, Neilson F, Hogarth H, et al. Cumulative attendance, assessment and cancer detection rate over four screening rounds in five English breast-screening programmes: a retrospective study. *J Public Health (Oxf)* 2007; 29: 275–80. Epub 2007 May 22.

34. Kobayashi E, Ueda Y, Matsuzaki S, et al. Biomarkers for screening, diagnosis, and monitoring of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2012; 21: 1902–12.

35. Hellquist BN, Duffy SW, Abdsaleh S, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. *Cancer* 2011; 117: 714–22.

36. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008; 246: 376–83. doi: 10.1148/radiol.2461070200.

Mini Review

Appendix: Literature Searches

Searches performed per database on 21–02-2017

- Embase

(mammography/exp OR 'breast tumor'/exp OR (mammogra* OR echomammogra* OR breast):ab,ti) AND (screening/de OR 'screening test'/de OR 'mass screening'/de OR 'cancer screening'/de OR (screen* OR (mammogra* NEAR/3 (routine* OR repeat*))):ab,ti) AND ('middle aged'/de OR premenopause/de OR ('middle aged' OR ((35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR fourt* OR thirt*) NEAR/3 (age* OR year*)) OR fourties* OR ((under OR below OR before OR younger OR 'less than') NEAR/3 (50 OR 49)) OR ((young* OR earl* OR start*) NEAR/3 age*) OR premenopaus* OR (pre NEXT/1 menopaus*)):ab,ti) AND ((random* OR factorial* OR crossover* OR (cross NEXT/1 over*) OR placebo* OR ((doubl* OR singl*) NEXT/1 blind*) OR assign* OR allocat* OR volunteer*):ab,ti OR 'crossover procedure'/de OR 'double-blind procedure'/de OR 'randomized controlled trial'/de OR 'single-blind procedure'/de) AND [english]/lim NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Conference Paper]/lim OR [Editorial]/lim)

- Medline (OvidSP)

(mammography/OR exp Breast Neoplasms/OR (mammogra* OR echomammogra* OR breast).ab,ti.) AND (Mass Screening/OR (screen* OR (mammogra* ADJ3 (routine* OR repeat*))).ab,ti.) AND (premenopause/OR (middle aged OR ((35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR fourt* OR thirt*) ADJ3 (age* OR year*)) OR fourties* OR (("under" OR "below" OR "before" OR "younger" OR "less than")

ADJ3 (50 OR 49)) OR ((young* OR earl* OR start*) ADJ3 age*) OR premenopaus* OR (pre ADJ menopaus*)).ab,ti.) AND (Clinical Trial.pt. OR randomized.ab,ti. OR placebo.ab,ti. OR dt.fs. OR randomly.ab,ti. OR trial.ab,ti. OR groups.ab,ti.) AND english.la. NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt.

- Cochrane Library

((mammogra* OR echomammogra* OR breast):ab,ti) AND ((screen* OR (mammogra* NEAR/3 (routine* OR repeat*))):ab,ti) AND (('middle aged' OR ((35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR fourt* OR thirt*) NEAR/3 (age* OR year*)) OR fourties* OR ((under OR below OR before OR younger OR 'less than') NEAR/3 (50 OR 49)) OR ((young* OR earl* OR start*) NEAR/3 age*) OR premenopaus* OR (pre NEXT/1 menopaus*)):ab,ti)

- PubMed

((mammogra*[tiab] OR echomammogra*[tiab] OR breast)) AND ((screen*[tiab] OR (mammogra*[tiab] AND (routine*[tiab] OR repeat*[tiab])))) AND (("middle aged"[tiab] OR ((35[tiab] OR 36[tiab] OR 37[tiab] OR 38[tiab] OR 39[tiab] OR 40[tiab] OR 41[tiab] OR 42[tiab] OR 43[tiab] OR 44[tiab] OR 45[tiab] OR 46[tiab] OR 47[tiab] OR 48[tiab] OR 49[tiab] OR fourt*[tiab] OR thirt*[tiab]) AND (age[tiab] OR aged[tiab] OR year*[tiab])) OR fourties*[tiab] OR (("under"[tiab] OR "below"[tiab] OR "before"[tiab] OR "younger"[tiab] OR "less than"[tiab]) AND (50[tiab] OR 49[tiab])) OR ((young*[tiab] OR earl*[tiab] OR start*[tiab]) AND (age[tiab] OR aged[tiab])) OR premenopaus*[tiab] OR (pre menopaus*[tiab]))) AND (randomized OR placebo OR randomly OR trial OR groups) AND english[la] AND publisher[sb]