EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Directorate F - Health, Consumers & Reference Materials (Ispra)
Health in Society

European Commission Initiative on Breast Cancer (ECIBC): European guidelines on breast cancer screening and diagnosis

European Commission, Via Enrico Fermi 2749, I-21027 Ispra (Varese) - Italy. Telephone: (39)0332-78-9111.
E-mail: JRC-CANCER-POLICY-SUPPORT@ec.europa.eu

QUESTION

Should organised mammography screening vs. no mammography screening be used for early detection of breast cancer in women aged 45 to 49?	
POPULATION:	Women aged 45 to 49
INTERVENTION:	organised mammography screening
COMPARISON:	no mammography screening
MAIN OUTCOMES:	Breast cancer mortality (short case accrual); Breast cancer mortality (longest case accrual available); Other cause mortality; Stage IIA breast cancer or higher; Stage III+ breast cancer or tumour size ≥40 mm; Rate of mastectomies; Provision of chemotherapy; Overdiagnosis (long case accrual); Quality of life (inferred from psychological effects); False-positive related adverse effects (psychological distress); and False-positive related adverse effects (biopsies and surgeries)
SETTING:	European Union
PERSPECTIVE:	Population (National Health System)
BACKGROUND:	Although mammography screening has both potential benefits and harms many countries have organised programmes for women aged 50 or older. However, there continues to be debate about recommendations for mammography screening, generally (Jorgensen 2009, Arie 2014), and particularly for women aged 40 to 49 (Petitti 2010).
CONFLICT OF INTEREST:	Management of Conflicts of Interests (CoI): CoIs for all Guidelines Development Group (GDG) members were assessed and managed by the Joint Research Center (JRC) following an established procedure in line with the European Commission rules. GDG member participation in the development of the recommendations was according to CoI disclosure. Consequently, for this particular question, the following GDG members were recused from voting: Roberto d'Amico, Jan Danes, Axel Gräwingholt and Ruben van Engen.

ASSESSMENT

Problem

Is the problem a priority?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	Breast cancer is the second most common cancer in the world and, by far, the most frequent cancer among women with an estimated 1.67 million new cancer cases diagnosed in 2012—accounting for 25% of all cancers (GLOBOCAN 2012). Breast cancer ranks as the fifth leading cause of cancer death worldwide and it is the second leading cause of cancer-related death in developed regions (GLOBOCAN 2012). In the European Union, 367 090 women were diagnosed of breast cancer and 92 000 women died from the disease in 2012 (Ferlay 2013). Breast cancer ranks fourth among the top five cancers with the highest disease burden (Tsilidis 2016). Annual incidence of breast cancer in the EU among women aged 45 to 49 is 1.7 per 1 000 and mortality is 0.2 per 1 000 per year (GLOBOCAN 2012)	The GDG prioritised this question for the ECIBC.

Desirable Effects

How substantial are the desirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Trivial ○ Small ● Moderate		These studies used an 'intention-to-treat' analysis thus, a per protocol approach would lead to larger

○ Large
○ Varies
○ Don't know

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)	
				Risk with no mammography screening	Risk difference with organised mammography screening
Breast cancer mortality (short case accrual) for women under 50 follow up: mean 16.8 years	348478 (8 RCTs)[1,2,3,4,5,6,7,a]	⊕⊕⊕◯ MODERATE[b,c,d]	**RR 0.88** (0.76 to 1.02)	Low	
				400 per 100.000[e]	**48 fewer per 100.000** (96 fewer to 8 more)
				High	
				700 per 100.000[f]	**84 fewer per 100.000** (168 fewer to 14 more)
Breast cancer mortality (longest case accrual available) for women under 50 follow up: mean 15.2 years	348478 (8 RCTs)[1,10,5,7,8,9,a]	⊕⊕⊕◯ MODERATE[b,c,d]	**RR 0.92** (0.83 to 1.02)	Low	
				480 per 100.000[e]	**38 fewer per 100.000** (82 fewer to 10 more)
Other cause mortality follow up: mean 10.8 years	290417 (6 RCTs)[11,12,13,14,8,a]	⊕◯◯◯ VERY LOW[b,c,d,g]	**RR 1.04** (0.95 to 1.15)	Low	
				2.500 per 100.000[e]	**100 more per 100.000** (125 fewer to 375 more)
Breast cancer stage IIA or higher follow up: mean 13.6 years[h]	300307 (5 RCTs)[10,12,15,16,4,7,9,a]	⊕◯◯◯ VERY LOW[d,i,j]	**RR 0.88** (0.78 to 0.99)	Low	
				380 per 100.000[e]	**46 fewer per 100.000** (84 fewer to 4 fewer)
Breast cancer stage - stage III+ or tumour size ≥40 mm follow up:	274194 (4 RCTs)[12,15,16,7,9,a]	⊕⊕◯◯ LOW[b,c,d]	**RR 0.98** (0.74 to 1.29)	Low	
				90 per	**2 fewer per 100.000**

absolute effects.

GDG members mentioned that modelling studies describing quality and duration of 'life gained' should be considered.

Long case accrual may dilute the effect of the intervention as for some trials it will include cases diagnosed after closure of the trial when both arms are receiving the same intervention. Therefore, we performed a sensitivity analysis including only studies that reported long case accrual estimates and we observed a small although not significant diluting effect (RR 0.92; 95% CI 0.83 to 1.02).

GDG members agreed that the desirable health effects differ by age at first screening. For women in the 45 to 49 age group, the GDG members agreed these women would have larger anticipated beneficial health effects (moderate effects) compared to women aged 40 to 44 due to higher absolute incidence and mortality from breast cancer in women aged 45-49 than in women aged 40-44. The percentage mortality reduction does not differ significantly from that observed in women aged 50 to 69; although there is substantial observational evidence for a benefit in women aged 45 to 49 (see evidence profile).

Test accuracy is poorer in younger women, largely due to mammographic breast density.

Digital mammography, which was not in use at the time of most of the studies reviewed here, may result in

mean 13.5 years[h]				100.000[e]	(23 fewer to 26 more)	
Rate of mastectomies	249550 (5 RCTs)[14,17,18,19,20,a]	⊕⊕◯◯ LOW[b,c,k]	**RR 1.20** (1.11 to 1.30)[l]	Low		
				900 per 100.000[e]	**180 more per 100.000** (99 more to 270 more)	
Provision of chemotherapy	99454 (2 RCTs)[14,19,20,a]	⊕◯◯◯ VERY LOW[c,d,k,m,n]	**RR 0.86** (0.53 to 1.40)[l]	Low		
				400 per 100.000[e]	**56 fewer per 100.000** (188 fewer to 160 more)	
Overdiagnosis (population perspective)	50430 (1 RCT)[12,a]	⊕⊕⊕◯ MODERATE[c]	-	12.4% (95% CI 9.9%-14.9%) [o]		
Overdiagnosis (woman perspective)	50430 (1 RCT)[12,a]	⊕⊕⊕◯ MODERATE[c]	-	22.7% (95% CI 18.4%-27.0%) [p]		
Quality of life (inferred from psychological effects)[h]	(54 observational studies)[21]	⊕⊕◯◯ LOW[q]	-	One systematic review with 54 studies included -no meta-analysis - (Brett 2005). Mammographic screening does not appear to create anxiety in women who are given a clear result after a mammogram and subsequently placed on routine recall. Mixed results about anxiety in women recalled for further testing: several studies reported transient or long term (from 6 months to 1 year after recall) anxiety, while other studies reported no differences in anxiety levels. The nature and extent of further testing seem to determine the extent of anxiety.		
False-positive related adverse effects (psychological distress)[h]	(24 observational studies)[22,23]	⊕⊕◯◯ LOW	-	Two systematic reviews. One review included 17 studies and found that women who received a false-positive mammogram result had greater distress, fear, anxiety, and worry about breast cancer (Saltz 2010). The second review included 7 studies, the psychological distress using diseases-specific measurements		

greater test accuracy in women aged 45 to 49.

In the Sweden Mammography Screening of Young Women (SCRY) cohort, which compared breast cancer mortality between women invited and not invited to screening; RRs of 0.82 (95% CI, 0.67-1.00) and 0.63 (95% CI, 0.54-0.75) for the age groups of 40 to 44 and 45 to 49 years were respectively reported. The weighted RR for the 40 to 49 years did not differ from the unweighted estimate of 0.71 (95% CI, 0.62-0.80).

				in women (age not specified) with a false-positive mammogram at 35 months after the last assessment was ; for women that needed further mammography RR=1.28 (95%CI 0.82-2.00); for women placed in early recall the RR=1.82 (95%CI 1.22-2.72); for women that needed a fine needle puncture aspiration RR=1.80 (95%CI 1.17-2.77); for women that needed a biopsy RR=2.07 (95%CI 1.22-3.52); no differences in generic measures of general anxiety and depression were observed at 6 weeks after assessment and 3 months after screening Bond (2013).
False-positive related adverse effects (biopsies and surgeries)[h]	(4 observational studies)[24]	⊕◯◯◯ VERY LOW[r]	-	Results from literature review (4 studies, 390 000 women aged 50 to 69) showed an overall false-positive screening result of 19.7% in women undergoing 10 biennial screening tests (pooled risk estimate based on 3 studies; range 8 - 21%). This was related to a 2.9% pooled cumulative risk of an invasive procedure with benign outcome (range 1.8% to 6.3%; based on 2 studies) and 0.9% risk of undergoing surgical intervention with benign outcome (based on 1 study) (Hofvind 2012). Cross-sectional data from the EUNICE Project (women aged 50 to 69): 17 countries, 20 screening programmes, 1.7 million initial screens, 5.9 million subsequent screens; showed that 2.2% and 1.1% of all screening examinations resulted in needle biopsy among women without breast cancer (initial and subsequent screens, respectively). In addition, 0.19% and 0.07% of all screening examinations resulted in surgical interventions among women without breast cancer (initial and subsequent screens, respectively).

1. Miller AB, Baines CJ,To T,Wall C.. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged

40 to 49 years.. CMAJ; 1992.
2. Tabar L, Duffy SW,Yen MF,Warwick J,Vitak B,Chen HH,Smith RA.. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point.. J Med Screen; 2002.
3. S, Shapiro. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan.. J Natl Cancer Inst Monogr; 1997.
4. Bjurstam NG, Björneld LM,Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer; 2016.
5. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet; 2002.
6. Moss SM, Cuckle H,Evans A,Johns L,Waller M,Bobrow L, Group., Trial,Management. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years´follow-up: a randomised controlled trial. Lancet Oncol; 2015.
7. Habbema JD, van Oortmarssen GJ,van Putten DJ,Lubbe JT,van der Maas PJ.. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York study.. J Natl Cancer Inst.; 1986.
8. Moss SM, Wale C,Smith R,Evans A,Cuckle H,Duffy SW.. 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.
9. Tabar L, Fagerberg G,Chen HH,Duffy SW,Smart CR,Gad A,et al.. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial.. Cancer. ; 1995.
10. Bjurstam N1, Björneld L,Warwick J,Sala E,Duffy SW,Nyström L,et al. The Gothenburg Breast Screening Trial.. Cancer; 2003.
11. Bjurstam N, Björneld L,Duffy SW,Smith TC,Cahlin E,Eriksson O,et al.. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization.. Cancer; 1997.
12. Miller AB, To T,Baines CJ,Wall C.. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years.. Ann Intern Med.; 2002 .
13. Tabar L, Fagerberg G,Duffy SW,and N E Day. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit.. J Epidemiol Community Health; 1989 .
14. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials.. Lancet.; 2002.
15. Chu KC, Smart CR,Tarone RE.. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial.. J Natl Cancer Inst; 1988.
16. Moss S, Waller M,Anderson TJ,Cuckle H. Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures.. Br J Cancer.; 2005.
17. AB, Miller. The costs and benefits of breast cancer screening.. Am J Prev Med; 1993.
18. J, Frisell. Mammographic screening for breast cancer [thesis]. Stockholm: Södersjukhuset; 1989.
19. Andersson I, Aspegren K,Janzon L, Landberg T,Lindholm K,Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial.. 1988; 1988.
20. Tabar L, Chen HH,Duffy SW,Krusemo UB.. Primary and adjuvant therapy, prognostic factors and survival in 1053 breast cancers diagnosed in a trial of mammography screening.. Jpn J Clin Oncol. ; 1999.
21. Brett J, Bankhead C,Henderson B,Watson E,Austoker J.. The psychological impact of mammographic screening. A systematic review.. Psychooncology; 2005.
22. Bond M, Pavey T,Welch K,Cooper C,Garside R,Dean S,et al.. Systematic review of the psychological consequences of false-positive screening mammograms.. Health Technol Assess; 2013.
23. Salz T, Richman AR,Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes..

Psychooncology.; 2010.
24. Hofvind S1, Ponti A,Patnick J,Ascunce N,Njor S,Broeders M,et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes.. J Med Screen. ; 2012.

a. The reference listed in the evidence profiles correspond to the specific publications used to extract crude data for estimating the outcomes´ effect sizes. Additional reference describing the characteristics of the included studies can be found in the document´s main text of this systematic review.
b. Some studies used methods that would not be accepted for random allocation today. One study had non-blinded assessment of 'cause of death'. The GDG felt that the CNBSS-1 possibly had issues with achieving prognostic balance. The GDG felt that lack of allocation concealment in this set of studies did not lead to high risk of bias. Given the lack of single trials driving the overall results and similarity in effect sizes (the test for subgroup differences - low vs high risk of bias trials - was non-significant) and overlapping confidence intervals (CIs), the risk of bias was rated as 'not serious'.
c. Trials were conducted more than 20 years ago. Currently, women have higher adherence to breast cancer screening and the quality control of screening and the care of breast cancer have improved. A large non-randomised study (Hellquist B 2011) showed a reduced risk for breast cancer deaths in women aged 40 to 49 years invited to screening, compared with women not invited (RR=0.74; 95%CI, 0.66-0.83) which is consistent with the results seen in the RCTs. The GDG did not downgrade for indirectness for breast cancer mortality but considered it serious for other outcomes.
d. 95% CI probably crosses the clinical decision threshold (as the CI is wide, a different clinical decision regarding the intervention may be taken depending on whether the lower or the higher limit is considered).
e. Median or mean of the control group of the included studies unless otherwise specified.
f. Baseline risk calculated from the ITACAN database. http://itacan.ispo.toscana.it/italian/itacan.htm
g. Unexplained inconsistency with statistical heterogeneity (I^2 = 62%, P = 0.02).
h. Importance of the outcome was lowered from 'critical' to 'important' because the GDG members felt this outcome influenced neither the direction nor the strength of the recommendation
i. Some studies were sub-optimally randomised and had non-blinded assessment of stage of disease; when analysis was restricted to low risk of bias trials, the risk estimate was non-significant.
j. Indirectness same as for women aged 50 to 69.
k. Population include women aged 40-74. Therefore, a much broader age range than the age group studied here. Observational studies do not confirm these results, instead they provide opposite results.
l. Due to lead time, there may be greater numbers of cancers to be treated in the screened group, during the period of observation, which may lead to overestimation of the rate of chemotherapy and mastectomies in the screened group
m. Unexplained inconsistency with statistical heterogeneity (I^2 = 71%, P = 0.06).
n. Chemotherapy protocols and indications have significantly changed (e.g. node status was not determined in earlier studies).
o. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed over whole follow-up period in women invited for screening (population perspective).
p. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed during screening period in women invited for screening (woman perspective).
q. Unexplained inconsistency for variability in anxiety in the group of women recalled for further testing.

r. Studies included women aged 50 to 69. Estimates for the 45-49 age stratum are likely to be higher.

Undesirable Effects

How substantial are the undesirable anticipated effects?

JUDGEMENT

○ Large
● Moderate
○ Small
○ Trivial
○ Varies
○ Don't know

RESEARCH EVIDENCE

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)	
				Risk with no mammography screening	Risk difference with organised mammography screening
Breast cancer mortality (short case accrual) for women under 50 follow up: mean 16.8 years	348478 (8 RCTs)[1,2,3,4,5,6,7,a]	⊕⊕⊕◯ MODERATE[b,c,d]	**RR 0.88** (0.76 to 1.02)	Low	
				400 per 100.000[e]	**48 fewer per 100.000** (96 fewer to 8 more)
				High	
				700 per 100.000[f]	**84 fewer per 100.000** (168 fewer to 14 more)
Breast cancer mortality (longest case accrual available) for women under 50 follow up: mean 15.2 years	348478 (8 RCTs)[1,10,5,7,8,9,a]	⊕⊕⊕◯ MODERATE[b,c,d]	**RR 0.92** (0.83 to 1.02)	Low	
				480 per 100.000[e]	**38 fewer per 100.000** (82 fewer to 10 more)
Other cause mortality follow up: mean 10.8 years	290417 (6 RCTs)[11,12,13,14,8,a]	⊕◯◯◯ VERY LOW[b,c,d,g]	**RR 1.04** (0.95 to 1.15)	Low	
				2.500 per 100.000[e]	**100 more per 100.000** (125 fewer to 375 more)

ADDITIONAL CONSIDERATIONS

Overdiagnosis and its magnitude are not greatly influenced by age at first screening.

Overdiagnosis estimates from both CNBSS1 and CNBSS2 may have been overestimated by subsequent screening in the population (both organised and opportunistic) after screening ceased in the CNBSS in 1988. Thus, while at 25 years of follow-up a non-statistically significant excess of all breast cancers was observed in the intervention arm of CNBSS trials (difference 2.6; 95%CI -0.8 to 5.9), the excess rate of in-situ/invasive breast cancers actually increased over the first-years post-screening in the CNBSS1, and dramatically decreased after the 10 years post-screening in the CNBSS2.

Due to lead time (diagnosis time being brought forward with screening), there may be greater numbers of cancers to be treated in the screened group, during the period of observation, which may lead to an increased rate of chemotherapy and mastectomies in the screened group.

False positive rates have been observed to be higher in women under the age of 50 than in women aged 50 to 69.

	Breast cancer stage IIA or higher follow up: mean 13.6 years[h]	300307 (5 RCTs)[10,12,15,16,4,7,9,a]	⊕◯◯◯ VERY LOW[d,i,j]	**RR 0.88** (0.78 to 0.99)	Low		The number of false positives will depend on the age at first screening. The GDG considered this effect to be moderate. Radiation risk is higher in younger women. The radiation exposure and associated risk is dependent on the screening method and frequency that, in turn, will influence the balance of benefits and harms.
					380 per 100.000[e]	**46 fewer per 100.000** (84 fewer to 4 fewer)	
	Breast cancer stage - stage III+ or tumour size ≥40 mm follow up: mean 13.5 years[h]	274194 (4 RCTs)[12,15,16,7,9,a]	⊕⊕◯◯ LOW[b,c,d]	**RR 0.98** (0.74 to 1.29)	Low		
					90 per 100.000[e]	**2 fewer per 100.000** (23 fewer to 26 more)	
	Rate of mastectomies	249550 (5 RCTs)[14,17,18,19,20,a]	⊕⊕◯◯ LOW[b,c,k]	**RR 1.20** (1.11 to 1.30)[l]	Low		
					900 per 100.000[e]	**180 more per 100.000** (99 more to 270 more)	
	Provision of chemotherapy	99454 (2 RCTs)[14,19,20,a]	⊕◯◯◯ VERY LOW[c,d,k,m,n]	**RR 0.86** (0.53 to 1.40)[l]	Low		
					400 per 100.000[e]	**56 fewer per 100.000** (188 fewer to 160 more)	
	Overdiagnosis (population perspective)	50430 (1 RCT)[12,a]	⊕⊕⊕◯ MODERATE[c]	-	12.4% (95% CI 9.9%-14.9%) [o]		
	Overdiagnosis (woman perspective)	50430 (1 RCT)[12,a]	⊕⊕⊕◯ MODERATE[c]	-	22.7% (95% CI 18.4%-27.0%) [p]		
	Quality of life (inferred from psychological effects)[h]	(54 observational studies)[21]	⊕⊕◯◯ LOW[q]	-	One systematic review with 54 studies included -no meta-analysis - (Brett 2005). Mammographic screening does not appear to create anxiety in women who are given a clear result after a mammogram and subsequently placed on routine recall. Mixed results about anxiety in women recalled for further testing: several studies reported transient or long term (from 6 months to 1 year after recall) anxiety, while other studies reported no differences in anxiety levels. The nature and	

				extent of further testing seem to determine the extent of anxiety.
False-positive related adverse effects (psychological distress)[h]	(24 observational studies)[22,23]	⊕⊕◯◯ LOW	-	Two systematic reviews. One review included 17 studies and found that women who received a false-positive mammogram result had greater distress, fear, anxiety, and worry about breast cancer (Saltz 2010). The second review included 7 studies, the psychological distress using diseases-specific measurements, in women (age not specified) with a false-positive mammogram at 35 months after the last assessment was ; for women that needed further mammography RR=1.28 (95%CI 0.82-2.00); for women placed in early recall the RR=1.82 (95%CI 1.22-2.72); for women that needed a fine needle puncture aspiration RR=1.80 (95%CI 1.17-2.77); for women that needed a biopsy RR=2.07 (95%CI 1.22-3.52); no differences in generic measures of general anxiety and depression were observed at 6 weeks after assessment and 3 months after screening Bond (2013).
False-positive related adverse effects (biopsies and surgeries)[h]	(4 observational studies)[24]	⊕◯◯◯ VERY LOW[r]	-	Results from literature review (4 studies, 390 000 women aged 50 to 69) showed an overall false-positive screening result of 19.7% in women undergoing 10 biennial screening tests (pooled risk estimate based on 3 studies; range 8 - 21%). This was related to a 2.9% pooled cumulative risk of an invasive procedure with benign outcome (range 1.8% to 6.3%; based on 2 studies) and 0.9% risk of undergoing surgical intervention with benign outcome (based on 1 study) (Hofvind 2012). Cross-sectional data from the EUNICE Project (women aged 50 to 69): 17 countries, 20 screening programmes, 1.7 million initial screens, 5.9 million subsequent screens; showed that 2.2% and 1.1% of all screening examinations resulted in needle

				biopsy among women without breast cancer (initial and subsequent screens, respectively). In addition, 0.19% and 0.07% of all screening examinations resulted in surgical interventions among women without breast cancer (initial and subsequent screens, respectively).

1. Miller AB, Baines CJ,To T,Wall C.. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years.. CMAJ; 1992.
2. Tabar L, Duffy SW,Yen MF,Warwick J,Vitak B,Chen HH,Smith RA.. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point.. J Med Screen; 2002.
3. S, Shapiro. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan.. J Natl Cancer Inst Monogr; 1997.
4. Bjurstam NG, Björneld LM,Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer; 2016.
5. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet; 2002.
6. Moss SM, Cuckle H,Evans A,Johns L,Waller M,Bobrow L, Group., Trial,Management. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years´follow-up: a randomised controlled trial. Lancet Oncol; 2015.
7. Habbema JD, van Oortmarssen GJ,van Putten DJ,Lubbe JT,van der Maas PJ.. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York study.. J Natl Cancer Inst.; 1986.
8. Moss SM, Wale C,Smith R,Evans A,Cuckle H,Duffy SW.. 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.
9. Tabar L, Fagerberg G,Chen HH,Duffy SW,Smart CR,Gad A,et al.. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial.. Cancer. ; 1995.
10. Bjurstam N1, Björneld L,Warwick J,Sala E,Duffy SW,Nyström L,et al. The Gothenburg Breast Screening Trial.. Cancer; 2003.
11. Bjurstam N, Björneld L,Duffy SW,Smith TC,Cahlin E,Eriksson O,et al.. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization.. Cancer; 1997.
12. Miller AB, To T,Baines CJ,Wall C.. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years.. Ann Intern Med.; 2002 .
13. Tabar L, Fagerberg G,Duffy SW,and N E Day. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit.. J Epidemiol Community Health; 1989 .
14. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials.. Lancet.; 2002.
15. Chu KC, Smart CR,Tarone RE.. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial.. J Natl Cancer Inst; 1988.
16. Moss S, Waller M,Anderson TJ,Cuckle H. Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures.. Br J Cancer.; 2005.
17. AB, Miller. The costs and benefits of breast cancer screening.. Am J Prev Med; 1993.

18. J, Frisell. Mammographic screening for breast cancer [thesis]. Stockholm: Södersjukhuset; 1989.
19. Andersson I, Aspegren K,Janzon L, Landberg T,Lindholm K,Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial.. 1988; 1988.
20. Tabar L, Chen HH,Duffy SW,Krusemo UB.. Primary and adjuvant therapy, prognostic factors and survival in 1053 breast cancers diagnosed in a trial of mammography screening.. Jpn J Clin Oncol. ; 1999.
21. Brett J, Bankhead C,Henderson B,Watson E,Austoker J.. The psychological impact of mammographic screening. A systematic review.. Psychooncology; 2005.
22. Bond M, Pavey T,Welch K,Cooper C,Garside R,Dean S,et al.. Systematic review of the psychological consequences of false-positive screening mammograms.. Health Technol Assess; 2013.
23. Salz T, Richman AR,Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes.. Psychooncology.; 2010.
24. Hofvind S1, Ponti A,Patnick J,Ascunce N,Njor S,Broeders M,et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes.. J Med Screen. ; 2012.

a. The reference listed in the evidence profiles correspond to the specific publications used to extract crude data for estimating the outcomes´ effect sizes. Additional reference describing the characteristics of the included studies can be found in the document´s main text of this systematic review.
b. Some studies used methods that would not be accepted for random allocation today. One study had non-blinded assessment of 'cause of death'. The GDG felt that the CNBSS-1 possibly had issues with achieving prognostic balance. The GDG felt that lack of allocation concealment in this set of studies did not lead to high risk of bias. Given the lack of single trials driving the overall results and similarity in effect sizes (the test for subgroup differences - low vs high risk of bias trials - was non-significant) and overlapping confidence intervals (CIs), the risk of bias was rated as 'not serious'.
c. Trials were conducted more than 20 years ago. Currently, women have higher adherence to breast cancer screening and the quality control of screening and the care of breast cancer have improved. A large non-randomised study (Hellquist B 2011) showed a reduced risk for breast cancer deaths in women aged 40 to 49 years invited to screening, compared with women not invited (RR=0.74; 95%CI, 0.66-0.83) which is consistent with the results seen in the RCTs. The GDG did not downgrade for indirectness for breast cancer mortality but considered it serious for other outcomes.
d. 95% CI probably crosses the clinical decision threshold (as the CI is wide, a different clinical decision regarding the intervention may be taken depending on whether the lower or the higher limit is considered).
e. Median or mean of the control group of the included studies unless otherwise specified.
f. Baseline risk calculated from the ITACAN database. http://itacan.ispo.toscana.it/italian/itacan.htm
g. Unexplained inconsistency with statistical heterogeneity (I^2 = 62%, P = 0.02).
h. Importance of the outcome was lowered from 'critical' to 'important' because the GDG members felt this outcome influenced neither the direction nor the strength of the recommendation
i. Some studies were sub-optimally randomised and had non-blinded assessment of stage of disease; when analysis was restricted to low risk of bias trials, the risk estimate was non-significant.
j. Indirectness same as for women aged 50 to 69.
k. Population include women aged 40-74. Therefore, a much broader age range than the age group studied here. Observational studies do not confirm these results, instead they provide opposite results.
l. Due to lead time, there may be greater numbers of cancers to be treated in the screened group, during the period of observation, which

	may lead to overestimation of the rate of chemotherapy and mastectomies in the screened group m. Unexplained inconsistency with statistical heterogeneity (I^2 = 71%, P = 0.06). n. Chemotherapy protocols and indications have significantly changed (e.g. node status was not determined in earlier studies). o. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed over whole follow-up period in women invited for screening (population perspective). p. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed during screening period in women invited for screening (woman perspective). q. Unexplained inconsistency for variability in anxiety in the group of women recalled for further testing. r. Studies included women aged 50 to 69. Estimates for the 45-49 age stratum are likely to be higher.	

Certainty of evidence

What is the overall certainty of the evidence of effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ○ Low ● Moderate ○ High ○ No included studies	The overall certainty (i.e. quality) of the evidence was considered moderate, as this was the lowest quality among the critical outcomes—namely, breast cancer mortality and overdiagnosis.	

Values

Is there important uncertainty about or variability in how much people value the main outcomes?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Important uncertainty or variability ● Possibly important uncertainty or variability ○ Probably no important uncertainty or variability ○ No important uncertainty or variability ○ No known undesirable outcomes	A systematic review shows that participants place a low value on the psychosocial and physical effects of false-positive results and overdiagnosis (JRC Technical Report PICO 10-11, contract FWC443094012015; available upon request). Women generally consider these undesirable effects acceptable (low confidence in evidence). However, these findings are of limited value mainly given the significant concerns regarding the adequacy of the information provided to women, in order to make an informed decision about participation. Also, acceptability of false positive results is based on studies of participants who have already received a false positive result. Their preferences may differ from the general population. Another finding is that breast cancer screening represents a significant burden for some participants due to the associated psychological distress and inconvenience (moderate confidence in evidence). Regarding breast cancer diagnosis, very limited data is available addressing people's views. One of the main themes identified in the literature is that people disvalue highly the anxiety caused by delays in receiving diagnostic results, or by a lack of understanding of the tests due to suboptimal communication with physicians (moderate confidence in evidence). Also, people have a higher overall preference towards more comfortable, brief diagnostic procedures (moderate confidence in evidence). (JRC Technical Report PICO 10-11, contract FWC443094012015; available upon request)	

Balance of effects

Does the balance between desirable and undesirable effects favor the intervention or the comparison?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ● Probably favors the intervention ○ Favors the intervention ○ Varies ○ Don't know	No research evidence was identified.	GDG members agreed that first screening at age 45 had moderate desirable health effects and moderate undesirable health effects; however, consensus was not reached regarding the balance between these two. Sixteen members voted that the balance probably favours the intervention; five members voted that the balance does not favour either the intervention or the comparison; and one voting member abstained.

Resources required

How large are the resource requirements (costs)?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	Differences in required resources for mammography screening versus no screening in women aged 40 to 49 in the studies analysed may be related to the inclusion or not of costs related to the screening process, diagnostic techniques, treatment and follow-up of diagnosed women (1) (2). Screening costs for a cohort of 10,000 women aged 47 to 49 years have been estimated to be £ 420,000 in the UK. The cost of diagnosis for positive results would be £ 70,000, and screening would lead to a saving of £ 17,000 in treatment costs (£480 per screen-detected cancer, calculated from the difference in treatment costs for the trial control and intervention arms), giving a net screening cost of £ 473,000 per 10,000 screened women (3.5% discount rate) (1) Based on the results of (2), the total cost of breast cancer diagnosis, treatment and death in the absence of screening was estimated to be € 1,161,008 per 1000 women aged 50 to 74 years, followed over their lifetime (3.5% discount rate). Biennial screening will cost € 1,298,065 per 1000 women (aged 50-74) screened and the reported costs per 1000 women aged 40 to 74 is € 1,467,598. Therefore, the estimated cost of screening 1000 women aged 40 to 49 years would be € 169,533.	Varies by screening interval and by country and by the presence of opportunistic screening. GDG members judged the cost to be at least moderate. However, substantial differences could be observed in European countries without population-based screening programmes or in those programmes with different screening policies. Estimates refer to organised

		screening programmes. Local/regional/country level resource/cost analyses exist or are required to estimate the cost for each setting.

Certainty of evidence of required resources

What is the certainty of the evidence of resource requirements (costs)?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ● Low ○ Moderate ○ High ○ No included studies	The certainty of the evidence of resource requirements is low due to the study design of the included studies which were modelling studies based on observational data. In addition, the following differences were observed: In (1) model parameters were based on data from a triennial screening while data from (2) corresponded to biennial screening. The studies reported costs of screening, diagnosis, and treatment. Based on their data, total costs per extension of one round of triennial screening would be £47 per woman in the UK (2006 value) which is similar to the €61.3 per one round of biennial screening in the Netherlands (2014 value). The formal assessment of the certainty in the evidence for cost and resources used was made using GRADE criteria and reported in the Evidence Profile (JRC Technical Report PICO 14-15, contract FWC443094012015; available upon request).	Both studies assessed the extension of their current population-based screening programmes. As previously stated, substantial differences could be observed in European countries without population-based screening programmes or in those programmes with different screening policies.

Cost effectiveness

Does the cost-effectiveness of the intervention favor the intervention or the comparison?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ○ Probably favors the intervention ○ Favors the intervention ● Varies ○ No included studies	Based on the evidence provided by (2), the extension of biennial mammography screening starting at age 40 appears to be cost-effective at a 'willingness-to-pay' of €20 000 per life year gained (LYG) with an incremental cost-effectiveness ratio (ICER) of €10 826 per LYG starting at age 40 instead of age 45. On the contrary, based on the evidence provided by (1), the extension of triennial mammography screening in women aged 47 to 49 does not appear to be cost-effective at a 'willingness-to-pay' of £20 000 per Quality Adjusted Life Years (QALY). The probability of being cost-effective at this threshold was low (29%). The ICER per QALY gained for triennial screening was £27 400.	Differences in the cost-effectiveness results could be explained by the differences in setting, policy of the screening programmes, outcomes measures and type of technology used. Whereas (2) reported the ICER per LYG in the Netherlands, (1) reported the ICER per QALY in the UK. The negative effects of false-positive results in the UK significantly reduced QALYs. (2) assessed digital mammography while (1) assessed screen-film mammography.

		GDG members considered cost-effectiveness to vary based on the opposing results from the modelling studies.

Equity
What would be the impact on health equity?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Reduced ○ Probably reduced ○ Probably no impact ○ Probably increased ○ Increased ○ Varies ● Don't know	No research evidence was identified.	A systematic review on this topic was not carried out. However, the utilisation of cancer screening services may largely depend on the availability of national public screening programmes; although European findings highlight that inequalities are larger in countries without population-based screening programmes (Palència, 2010).

Acceptability
Is the intervention acceptable to key stakeholders?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	A systematic review (JRC Technical Report PICO 16-17, contract FWC443094032016; available upon request) found the following barriers associated with breast cancer screening: (*a*) lack of knowledge and misperceptions regarding preventive medicine and breast health (high confidence in evidence), (*b*) poor communication skills of healthcare providers (high confidence in evidence), (*c*) poor accessibility to breast screening, especially among women with disabilities (high confidence in evidence), (*d*) fear and stress related to the procedure and the possibility of cancer diagnosis (high confidence in evidence), (*e*) pain and discomfort during the procedure (moderate confidence in evidence), (*f*) embarrassment and shyness during the procedure (moderate confidence in evidence), (*g*) lack of support and encouragement from family members, caregivers and social network (moderate confidence in evidence), (*h*) lack of information regarding the available resources (low confidence in evidence) and (*i*) low prioritisation of breast cancer screening (low confidence in evidence).	Some GDG members described that some professional groups may find a screening programme not acceptable due to their financial interests.

Feasibility
Is the intervention feasible to implement?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	No research evidence was identified.	A systematic review on this topic was not carried out. Some countries do not have screening programmes mainly due to lack of resources and also infrastructure.

		Given that this recommendation would be additive to screening in older age groups (50 to 69), it was judged as being probably feasible to implement.

SUMMARY OF JUDGEMENTS

	JUDGEMENT						
PROBLEM	No	Probably no	Probably yes	**Yes**		Varies	Don't know
DESIRABLE EFFECTS	Trivial	Small	**Moderate**	Large		Varies	Don't know
UNDESIRABLE EFFECTS	Large	**Moderate**	Small	Trivial		Varies	Don't know
CERTAINTY OF EVIDENCE	Very low	Low	**Moderate**	High			No included studies
VALUES	Important uncertainty or variability	**Possibly important uncertainty or variability**	Probably no important uncertainty or variability	No important uncertainty or variability			No known undesirable outcomes
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	**Probably favors the intervention**	Favors the intervention	Varies	Don't know
RESOURCES REQUIRED	Large costs	**Moderate costs**	Negligible costs and savings	Moderate savings	Large savings	Varies	Don't know
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	**Low**	Moderate	High			No included studies
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	**Varies**	No included studies
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased	Varies	**Don't know**
ACCEPTABILITY	No	Probably no	Probably yes	**Yes**		Varies	Don't know
FEASIBILITY	No	Probably no	**Probably yes**	Yes		Varies	Don't know

TYPE OF RECOMMENDATION

Strong recommendation against the intervention	Conditional recommendation against the intervention	Conditional recommendation for either the intervention or the comparison	**Conditional recommendation for the intervention**	Strong recommendation for the intervention
○	○	○	●	○

CONCLUSIONS

Recommendation

For asymptomatic women aged 45 to 49 with an average risk of breast cancer, the ECIBC's Guideline Development Group (GDG) suggests mammography screening over no mammography screening, in the context of an organised screening programme (conditional recommendation, moderate certainty of the evidence).

Justification

Overall justification

The conditional recommendation in favour of mammography screening over no mammography screening, in the context of an organised screening programme, was a result of a balance of health effects that probably favours mammography screening, despite only moderate certainty in the evidence about these effects. GDG members agreed these women would have larger anticipated beneficial health effects (moderate effects) compared to women aged 40 to 44 due to higher absolute incidence and mortality from breast cancer in women aged 45-49 than in women aged 40-44 together with observational evidence showing a greater benefit in this age group (Hellquist 2011).

As agreement within the GDG for the direction of this recommendation could not be reached, voting among members without CoI took place: 17 members voted that it should be a conditional recommendation in favour of the intervention; 1 member voted that it should be a conditional recommendation against the intervention; 4 members abstained.

Detailed justification

Desirable Effects

Mammography, compared to no screening, did not significantly reduce the risk of breast cancer mortality (77 fewer breast cancer deaths per 100 000, with a range from 7 more to 147 fewer deaths, or 44 fewer breast cancer deaths per 100 000, with a range from 4 more to 84 fewer breast cancer deaths, using a 0.7% and 0.4% baseline risk, respectively) in women invited to screening over 16.4 years of follow-up (moderate quality evidence). Although there is substantial observational evidence for a benefit in women aged 45 to 49. Mammography, compared to no screening, reduced the risk of stage IIA breast cancer or higher (46 fewer cases of breast cancer per 100 000 women during mean 13.6 years of follow-up) (very low quality evidence) but did not reduce the risk of all cause mortality (low quality evidence), other cause mortality (very low quality evidence) and stage III+ breast cancer or tumour size ≥ 40 mm (low quality evidence).

Undesirable Effects

Women aged 40 to 74 randomised to 'invitation to screening' were more likely to undergo mastectomy (180 more mastectomies per 100 000 women) (low quality evidence). Overdiagnosis is estimated to be 12.4% (moderate quality evidence) from a population perspective and 22.7% from the perspective of a woman invited to screening (moderate quality evidence). The number of false-positives will depend on the age at first screening. Estimated cumulative risk of a false-positive screening result in women aged 50 to 69 undergoing 10 biennial screening tests was 19.7%. However, false-positive rates have been observed to be higher in women under age 50 than in women aged 50 to 69. In addition, 2.2% of women had a needle biopsy after the initial screening mammogram. False-positive mammograms are also associated with greater anxiety and distress about breast cancer as well as negative psychological consequences that may last up to three years (low quality evidence). Mammography screening compared with no screening did not increase the number of women aged 43 to 74 treated with chemotherapy (very low quality evidence). Women who had further testing following their routine mammogram experienced significant short-term anxiety.

Certainty of evidence

The overall certainty (i.e. quality) of the evidence was considered moderate, as this was the lowest quality (corresponding to the quality of other cause mortality) among the outcomes considered to be critical (breast cancer mortality and overdiagnosis).

Subgroup considerations

This recommendation does not apply to high-risk women (see recommendations for women with high breast density).

Implementation considerations

GDG members agreed on the need for additional imaging techniques in this age group, together with the need for shared decision making. Implementation in this age group should be done in such a way to allow further quantification of benefits and harms.

Monitoring and evaluation

Future monitoring and evaluation of screening services should consider risks and benefits in the context of evolving treatment and management protocols.

Monitoring and evaluation criteria are being developed within the ECIBC initiative.

Research priorities

1. Carry out evaluations of the efficacy of the intervention, time intervals, risk factors and stratification of women, as well as context specific cost-effectiveness in this age group.

2. Carry out studies addressing the role of other screening modalities (e.g. MRI) in this population.

REFERENCES SUMMARY

1. Madan J, Rawdin A,Stevenson M,Tappenden P.. A Rapid Response Economic Evaluation of the UK NHS Cancer Reform Strategy Breast Cancer Screening Program Extension via a Plausible Bounds Approach. Value Health; 2010.
2. Sankatsing VD, Heijnsdijk EA,van Luijt PA,van Ravesteyn NT,Fracheb

Evidence profile

Healthcare question	Should organised mammography screening vs. no mammography screening be used for early detection of breast cancer in women aged 45 to 49?
Date	April 2016
Authors	ECIBC Guidelines Development Group (GDG): Mariangela Autelitano, Bettina Borisch, Mireille Broeders, Xavier Castells, Roberto D'Amico, Edoardo Colzani, Jan Daneš, Chris De Wolf, Stephen Duffy, Patricia Fitzpatrick, Markus Follmann, Livia Giordano, Paolo Giorgi Rossi, Axel Gräwingholt, Solveig Hofvind, Lydia Ioannidou-Mouzaka, Susan Knox, Annette Lebeau, Helen Mcgarrigle, Lennarth Nyström, Elsa Pérez Gómez, Cecily Quinn, Peter Rabe, Holger Schünemann, Alberto Torresin, Ruben Van Engen, Cary Van Landsveld-Verhoeven, Sue Warman, Kenneth Young. Systematic Review team: Diogenes Seraphim, Pablo Alonso-Coello, Ivan Solà, Monica Ballesteros, Margarita Posso, Nadia Montero, Carlos Canelo. JRC Healthcare Quality team: Zuleika Saz-Parkinson, Donata Lerda.
Abbreviations	**CI:** Confidence interval **RR:** Risk ratio

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Organised mammography screening	No mammography screening	Relative (95% CI)	Absolute (95% CI)		
Breast cancer mortality (short case accrual) for women under 50 (follow up: mean 16.8 years)												
8 [1,2,3,4,5,6,7,a]	randomised trials	not serious [b]	not serious	not serious [c]	serious [d]	none	428/152344 (0.3%)	0.4% [e]	**RR 0.88** (0.76 to 1.02)	**48 fewer per 100,000** (from 96 fewer to 8 more) [g]	⊕⊕⊕◯ MODERATE	CRITICAL
								0.7% [f]		**84 fewer per 100,000** (from 168 fewer to 14 more)		
Breast cancer mortality (longest case accrual available) for women under 50 (follow up: mean 15.2 years)												
8 [1,5,7,8,9,10,a]	randomised trials	not serious [b]	not serious	not serious [c]	serious [d]	none	736/152344 (0.5%)	0.5% [e]	**RR 0.92** (0.83 to 1.02)	**38 fewer per 100,000** (from 82 fewer to 10 more)	⊕⊕⊕◯ MODERATE	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Organised mammography screening	No mammography screening	Relative (95% CI)	Absolute (95% CI)		
Other cause mortality (follow up: mean 10.8 years)												
6 [8,11,12,13,14,a]	randomised trials	not serious [b]	serious [h]	serious [c]	serious [d]	none	3349/120225 (2.8%)	2.5% [e]	**RR 1.04** (0.95 to 1.15)	**100 more per 100,000** (from 125 fewer to 375 more)	⊕◯◯◯ VERY LOW	IMPORTANT
Breast cancer stage IIA or higher (follow up: mean 13.6 years)[i]												
5 [4,7,9,10,12,15,16,a]	randomised trials	serious [j]	not serious	serious [k]	serious [d]	none	475/124473 (0.4%)	0.4% [e]	**RR 0.88** (0.78 to 0.99)	**46 fewer per 100,000** (from 84 fewer to 4 fewer)	⊕◯◯◯ VERY LOW	IMPORTANT
Breast cancer stage - stage III+ or tumour size ≥40 mm (follow up: mean 13.5 years)[i]												
4 [7,9,12,15,16,a]	randomised trials	not serious [b]	not serious	serious [c]	serious [d]	none	93/112681 (0.1%)	0.1% [e]	**RR 0.98** (0.74 to 1.29)	**2 fewer per 100,000** (from 23 fewer to 26 more)	⊕⊕◯◯ LOW	IMPORTANT
Rate of mastectomies												
5 [14,17,18,19,20,a]	randomised trials	not serious [b]	not serious	very serious [c,l]	not serious	none	1542/144920 (1.1%)	0.9% [e]	**RR 1.20** (1.11 to 1.30) [m]	**180 more per 100,000** (from 99 more to 270 more)	⊕⊕◯◯ LOW	IMPORTANT
Provision of chemotherapy												
2 [14,19,20,a]	randomised trials	not serious	serious [n]	very serious [c,l,o]	serious [d]	none	252/59677 (0.4%)	0.4% [e]	**RR 0.86** (0.53 to 1.40) [m]	**56 fewer per 100,000** (from 188 fewer to 160 more)	⊕◯◯◯ VERY LOW	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Organised mammography screening	No mammography screening	Relative (95% CI)	Absolute (95% CI)		
Overdiagnosis (population perspective)												
1 [12,a]	randomised trials	not serious	not serious	serious [c]	not serious	none	12.4% (95% CI 9.9%-14.9%) [p]				⊕⊕⊕◯ MODERATE	CRITICAL
Overdiagnosis (woman perspective)												
1 [12,a]	randomised trials	not serious	not serious	serious [c]	not serious	none	22.7% (95% CI 18.4%-27.0%) [q]				⊕⊕⊕◯ MODERATE	CRITICAL
Quality of life (inferred from psychological effects)[i]												
54 [21]	observational studies	not serious	not serious [r]	not serious	not serious	none	One systematic review with 54 studies included -no meta-analysis - (Brett 2005). Mammographic screening does not appear to create anxiety in women who are given a clear result after a mammogram and subsequently placed on routine recall. Mixed results about anxiety in women recalled for further testing: several studies reported transient or long term (from 6 months to 1 year after recall) anxiety, while other studies reported no differences in anxiety levels. The nature and extent of further testing seem to determine the extent of anxiety.				⊕⊕◯◯ LOW	IMPORTANT
False-positive related adverse effects (psychological distress)[i]												
24 [22,23]	observational studies	not serious	not serious	not serious	not serious	none	Two systematic reviews. One review included 17 studies and found that women who received a false-positive mammogram result had greater distress, fear, anxiety, and worry about breast cancer (Saltz 2010). The second review included 7 studies, the psychological distress using diseases-specific measurements, in women (age not specified) with a false-positive mammogram at 35 months after the last assessment was ; for women that needed further mammography RR=1.28 (95%CI 0.82-2.00); for women placed in early recall the RR=1.82 (95%CI 1.22-2.72); for women that needed a fine needle puncture aspiration RR=1.80 (95%CI 1.17-2.77); for women that needed a biopsy RR=2.07 (95%CI 1.22-3.52); no differences in generic measures of general anxiety and depression were observed at 6 weeks after assessment and 3 months after screening Bond (2013).				⊕⊕◯◯ LOW	IMPORTANT
False-positive related adverse effects (biopsies and surgeries)[i]												
4 [24]	observational studies	not serious	not serious	serious [s]	not serious	none	Results from literature review (4 studies, 390 000 women aged 50 to 69) showed an overall false-positive screening result of 19.7% in women undergoing 10 biennial screening tests (pooled risk estimate based on 3 studies; range 8 - 21%). This was related to a 2.9% pooled cumulative risk of an invasive procedure with benign outcome (range 1.8% to 6.3%; based on 2 studies) and 0.9% risk of undergoing surgical intervention with benign outcome (based on 1 study) (Hofvind 2012). Cross-sectional				⊕◯◯◯ VERY LOW	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Organised mammography screening	No mammography screening	Relative (95% CI)	Absolute (95% CI)		
							data from the EUNICE Project (women aged 50 to 69): 17 countries, 20 screening programmes, 1.7 million initial screens, 5.9 million subsequent screens; showed that 2.2% and 1.1% of all screening examinations resulted in needle biopsy among women without breast cancer (initial and subsequent screens, respectively). In addition, 0.19% and 0.07% of all screening examinations resulted in surgical interventions among women without breast cancer (initial and subsequent screens, respectively).				

Explanations

a. The reference listed in the evidence profiles correspond to the specific publications used to extract crude data for estimating the outcomes´ effect sizes. Additional reference describing the characteristics of the included studies can be found in the document´s main text of this systematic review.

b. Some studies used methods that would not be accepted for random allocation today. One study had non-blinded assessment of 'cause of death'. The GDG felt that the CNBSS-1 possibly had issues with achieving prognostic balance. The GDG felt that lack of allocation concealment in this set of studies did not lead to high risk of bias. Given the lack of single trials driving the overall results and similarity in effect sizes (the test for subgroup differences - low vs high risk of bias trials - was non-significant) and overlapping confidence intervals (CIs), the risk of bias was rated as 'not serious'.

c. Trials were conducted more than 20 years ago. Currently, women have higher adherence to breast cancer screening and the quality control of screening and the care of breast cancer have improved. A large non-randomised study (Hellquist B 2011) showed a reduced risk for breast cancer deaths in women aged 40 to 49 years invited to screening, compared with women not invited (RR=0.74; 95%CI, 0.66-0.83) which is consistent with the results seen in the RCTs. The GDG did not downgrade for indirectness for breast cancer mortality but considered it serious for other outcomes.

d. 95% CI probably crosses the clinical decision threshold (as the CI is wide, a different clinical decision regarding the intervention may be taken depending on whether the lower or the higher limit is considered).

e. Median or mean of the control group of the included studies unless otherwise specified.

f. Baseline risk calculated from the ITACAN database. http://itacan.ispo.toscana.it/italian/itacan.htm

g. A large large non-randomised study (Hellquist 2011) showed a reduced risk for breast cancer deaths in women aged 40 to 49 years invited to screening, compared with women not invited (RR 0.74; 95%CI, 0.66 to 0.83) which is consistent with the results seen in the RCTs.

h. Unexplained inconsistency with statistical heterogeneity (I^2 = 62%, P = 0.02).

i. Importance of the outcome was lowered from 'critical' to 'important' because the GDG members felt this outcome influenced neither the direction nor the strength of the recommendation

j. Some studies were sub-optimally randomised and had non-blinded assessment of stage of disease; when analysis was restricted to low risk of bias trials, the risk estimate was non-significant.

k. Indirectness same as for women aged 50 to 69.

l. Population include women aged 40-74. Therefore, a much broader age range than the age group studied here. Observational studies do not confirm these results, instead they provide opposite results.

m. Due to lead time, there may be greater numbers of cancers to be treated in the screened group, during the period of observation, which may lead to overestimation of the rate of chemotherapy and mastectomies in the screened group

n. Unexplained inconsistency with statistical heterogeneity (I^2 = 71%, P = 0.06).

o. Chemotherapy protocols and indications have significantly changed (e.g. node status was not determined in earlier studies).

p. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed over whole follow-up period in women invited for screening (population perspective).

q. Overdiagnosis calculated from CNBSS-1 trial, in which women in the control group were not offered mammography screening at the end of the trial. Excess cancers as a proportion of cancers diagnosed during screening period in women invited for screening (woman perspective).
r. Unexplained inconsistency for variability in anxiety in the group of women recalled for further testing.
s. Studies included women aged 50 to 69. Estimates for the 45-49 age stratum are likely to be higher.

References

1. Miller AB, Baines CJ,To T,Wall C.. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years.. CMAJ; 1992.
2. Tabar L, Duffy SW,Yen MF,Warwick J,Vitak B,Chen HH,Smith RA.. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point.. J Med Screen; 2002.
3. S, Shapiro. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan.. J Natl Cancer Inst Monogr; 1997.
4. Bjurstam NG, Björneld LM,Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer; 2016.
5. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet; 2002.
6. Moss SM, Cuckle H,Evans A,Johns L,Waller M,Bobrow L, Group., Trial,Management. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years ´follow-up: a randomised controlled trial. Lancet Oncol; 2015.
7. Habbema JD, van Oortmarssen GJ,van Putten DJ,Lubbe JT,van der Maas PJ.. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York study.. J Natl Cancer Inst.; 1986.
8. Moss SM, Wale C,Smith R,Evans A,Cuckle H,Duffy SW.. 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.
9. Tabar L, Fagerberg G,Chen HH,Duffy SW,Smart CR,Gad A,et al.. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial.. Cancer. ; 1995.
10. Bjurstam N1, Björneld L,Warwick J,Sala E,Duffy SW,Nyström L,et al. The Gothenburg Breast Screening Trial.. Cancer; 2003.
11. Bjurstam N, Björneld L,Duffy SW,Smith TC,Cahlin E,Eriksson O,et al.. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization.. Cancer; 1997.
12. Miller AB, To T,Baines CJ,Wall C.. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years.. Ann Intern Med.; 2002 .
13. Tabar L, Fagerberg G,Duffy SW,and N E Day. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit.. J Epidemiol Community Health; 1989 .
14. Nyström L, Andersson I,Bjurstam N,Frisell J,Nordenskjöld B,Rutqvist LE.. Long-term effects of mammography screening: updated overview of the Swedish randomised trials.. Lancet.; 2002.
15. Chu KC, Smart CR,Tarone RE.. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial.. J Natl Cancer Inst; 1988.
16. Moss S, Waller M,Anderson TJ,Cuckle H. Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures.. Br J Cancer.; 2005.
17. AB, Miller. The costs and benefits of breast cancer screening.. Am J Prev Med; 1993.
18. J, Frisell. Mammographic screening for breast cancer [thesis]. Stockholm: Södersjukhuset; 1989.
19. Andersson I, Aspegren K,Janzon L, Landberg T,Lindholm K,Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial.. 1988; 1988.
20. Tabar L, Chen HH,Duffy SW,Krusemo UB.. Primary and adjuvant therapy, prognostic factors and survival in 1053 breast cancers diagnosed in a trial of mammography screening.. Jpn J Clin Oncol. ; 1999.
21. Brett J, Bankhead C,Henderson B,Watson E,Austoker J.. The psychological impact of mammographic screening. A systematic review.. Psychooncology; 2005.

22. Bond M, Pavey T,Welch K,Cooper C,Garside R,Dean S,et al.. Systematic review of the psychological consequences of false-positive screening mammograms.. Health Technol Assess; 2013.
23. Salz T, Richman AR,Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes.. Psychooncology.; 2010.
24. Hofvind S1, Ponti A,Patnick J,Ascunce N,Njor S,Broeders M,et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes.. J Med Screen. ; 2012.