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 of 50 to 69?	
POPULATION:	Women aged of 50 to 69
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. A reassessment of the evidence for screening women aged 50 to 69 is appropriate considering advances in diagnosis and treatment of breast cancer.
CONFLICT OF INTEREST:	Management of Conflicts of Interests (CoI): CoIs of all Guidelines Development Group (GDG) members were assessed and managed by the Joint Research Centre (JRC) following an established procedure in line with European Commission rules. GDG member participation in the development of the recommendations was restricted, according to CoI disclosure. Consequently, for this particular question, the following GDG members were recused from voting: Mireille Broeders, Roberto d'Amico, Jan Danes, Patricia Fitzpatrick, Axel Gräwingholt, Elsa Pérez Gómez, Ruben van Engen, Cary van Landsveld-Verhoeven, and Kenneth Young.

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 for25% of all cancers (GLOBOCAN 2012). Breast cancer ranks as the fifth leading cause of cancer death worldwide and the second leading cause of cancer-related death in developed regions (citation). In the European Union, 367 090 women were diagnosed with breast cancer and 92 000 women died from the disease in 2012 (1). Breast cancer ranks fourth among the top five cancers with the highest disease burden (2). Annual incidence of breast cancer in the EU among women aged 50 to 69 is 2.7 per 1 000 and mortality is 0.5 per 1 000 (GLOBOCAN 2012)	The GDG prioritised this question for the ECIBC.

Desirable Effects

How substantial are the desirable anticipated effects?

JUDGEMENT

○ Trivial
○ Small
○ Moderate
● Large
○ 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) follow up: mean 17.6 years	249930 (6 RCTs)[1,2,3,4,5,6,a,b]	⨁⨁⨁⨁ HIGH[c,d,e]	**RR 0.77** (0.66 to 0.90)	Low	
				600 per 100.000[b]	**138 fewer per 100.000** (204 fewer to 60 fewer)
				Moderate	
				1.000 per 100.000	**230 fewer per 100.000** (340 fewer to 100 fewer)
				High	
				2.100 per 100.000	**483 fewer per 100.000** (714 fewer to 210 fewer)
Breast cancer mortality (longest case accrual) follow up: mean 15.5 years	249930 (6 RCTs)[3,5,6,7,8,a]	⨁⨁⨁⨁ HIGH[c,d,e]	**RR 0.77** (0.67 to 0.88)	Low	
				760 per 100.000[f]	**175 fewer per 100.000** (251 fewer to 91 fewer)
Breast cancer stage IIA or higher[g]	297965 (4 RCTs)[1,10,6,7,8,9,a]	⨁◯◯◯ VERY LOW[c,h,i,j]	**RR 0.80** (0.64 to 1.00)	Moderate	
				700 per 100.000[f]	**140 fewer per 100.000** (252 fewer to 0 fewer)
Breast cancer	170032	⨁⨁◯◯	**RR 0.62**	Low	

ADDITIONAL CONSIDERATIONS

These studies used an 'intention-to-treat' analysis thus, a per protocol approach would lead to even larger absolute effects.

Estimates from observational studies were similar to those described here (see evidence profile).

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, but the pooled result was similar to the one obtained for the short accrual estimates.

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.

As there was disagreement among GDG members regarding whether the effects were large or moderate, voting took place among the 18 GDG members: 15 GDG members voted that the effects were large. Two GDG members voted that the effects were moderate. One GDG member abstained. Eight members could not vote due to CoI.

stage III+ or tumour size ≥40 mm [g]	(3 RCTs)[10,6,8,9,a]	LOW[c,h]	(0.48 to 0.80)	170 per 100.000[f]	**65 fewer per 100.000** (88 fewer to 34 fewer)
Other cause mortality follow up: mean 9.6 years	119083 (3 RCTs)[11,9,a]	⊕⊕◯◯ LOW[c,j]	**RR 0.99** (0.95 to 1.04)	Low	
				6.600 per 100.000[f]	**66 fewer per 100.000** (330 fewer to 264 more)
Overdiagnosis (population perspective)	64117 (2 RCTs)[12,9,a]	⊕⊕⊕◯ MODERATE[c]	-	10.1% (95% CI 8.6%-11.6%) [k]	
Overdiagnosis (patient perspective)	64117 (2 RCTs)[12,9,a]	⊕⊕⊕◯ MODERATE[c]	-	17.3% (95%CI 14.7%-20.0%) [l]	
Rate of mastectomies[g]	249550 (5 RCTs)[13,14,15,16,3,a]	⊕⊕◯◯ LOW[c,m]	**RR 1.20** (1.11 to 1.30)[n]	Low	
				900 per 100.000[f]	**180 more per 100.000** (99 more to 270 more)
Provision of chemotherapy[g]	99454 (2 RCTs)[15,16,3,a]	⊕◯◯◯ VERY LOW[c,j,o,p]	**RR 0.86** (0.52 to 1.41)[n]	Low	
				400 per 100.000[f]	**56 fewer per 100.000** (192 fewer to 164 more)
Quality of life (inferred from psychological effects)[g]	(54 observational studies)[17]	⊕⊕◯◯ 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)[g]	(24 observational studies)[18,19]	⊕⊕◯◯ 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)[g]	(4 observational studies)[20]	⊕⊕◯◯ LOW	-	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. Bjurstam NG, Björneld LM,Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer; 2016.
2. S, Shapiro. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan.. J Natl Cancer Inst Monogr; 1997.
3. 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.
4. 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.
5. Miller AB, Baines CJ,To T,Wall C.. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years.. CMAJ.; 1992.
6. 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.
7. Bjurstam N1, Björneld L,Warwick J,Sala E,Duffy SW,Nyström L,et al. The Gothenburg Breast Screening Trial.. Cancer; 2003.
8. 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.
9. Miller AB, To T,Baines CJ,Wall C.. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years.. J Natl Cancer Inst.; 2000.
10. 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.
11. 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 .
12. Zackrisson S, Andersson I, Janzon L,Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ; 2006.
13. J, Frisell. Mammographic screening for breast cancer [thesis]. Stockholm: Södersjukhuset; 1989.
14. AB, Miller. The costs and benefits of breast cancer screening.. Am J Prev Med; 1993.
15. 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.
16. 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.
17. Brett J, Bankhead C,Henderson B,Watson E,Austoker J.. The psychological impact of mammographic screening. A systematic review.. Psychooncology; 2005.
18. 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.
19. Salz T, Richman AR,Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes.. Psychooncology.; 2010.
20. 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. The GDG felt that baseline risks higher than 0.6% should be considered to evaluate absolute effects (Breast Cancer Screening, IARC Handbook of Cancer Prevention Volume 15)
c. Trials were conducted more than 20 years ago. Currently, women have higher adherence to breast cancer screening while quality control of screening and breast cancer care have improved.
d. Despite concerns about indirectness from the trials, including the fact that the population age range of 40 to 74 is broader than the age range in this question, after considering evidence from contemporary non-randomised studies (Broeders 2012) the GDG decided not to downgrade the quality of evidence for indirectness.
e. 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-2 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 that 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'.
f. Median or mean of the control group of the included studies unless otherwise specified.
g. Importance of the outcome was lowered from 'critical' to 'important' because the members felt this outcome influenced neither the direction nor the strength of the recommendation.
h. Non-blinded assessment of breast cancer stage is a serious concern. GDG members decided to downgrade to 'serious' for risk of bias.
i. Unexplained inconsistency with statistical heterogeneity (I^2 = 70%, P = 0.02). While one study shows clear benefit, in three studies the 95%CI does not exclude important benefit or harm.
j. 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).
k. Estimate from a meta-analysis of 2 trials (CNBSS-2 and Malmo I) 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).
l. Estimate from a meta-analysis of 2 trials (CNBSS-2 and Malmo I) 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).
m. Observational studies do not confirm these results, instead they provide opposite results.
n. 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 an increased rate of chemotherapy and mastectomies in the screened group
o. Unexplained inconsistency with statistical heterogeneity (I^2 = 71%, P=0.06).
p. Chemotherapy protocols and indications have significantly changed (e.g. node status was not determined in earlier studies).
q. Unexplained inconsistency for variability in anxiety in the group of women recalled for further testing.

Undesirable Effects

How substantial are the undesirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Large ● Moderate ○ Small ○ Trivial ○ Varies ○ Don't know		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.

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) follow up: mean 17.6 years	249930 (6 RCTs)[1,2,3,4,5,6,a,b]	⨁⨁⨁⨁ HIGH[c,d,e]	**RR 0.77** (0.66 to 0.90)	Low	
				600 per 100.000[b]	**138 fewer per 100.000** (204 fewer to 60 fewer)
				Moderate	
				1.000 per 100.000	**230 fewer per 100.000** (340 fewer to 100 fewer)
				High	
				2.100 per 100.000	**483 fewer per 100.000** (714 fewer to 210 fewer)
Breast cancer mortality (longest case accrual) follow up: mean 15.5 years	249930 (6 RCTs)[3,5,6,7,8,a]	⨁⨁⨁⨁ HIGH[c,d,e]	**RR 0.77** (0.67 to 0.88)	Low	
				760 per 100.000[f]	**175 fewer per 100.000** (251 fewer to 91 fewer)
Breast cancer stage IIA or higher[g]	297965 (4 RCTs)[1,10,6,7,8,9,a]	⨁◯◯◯ VERY LOW[c,h,i,j]	**RR 0.80** (0.64 to 1.00)	Moderate	
				700 per 100.000[f]	**140 fewer per 100.000** (252 fewer to 0 fewer)
Breast cancer stage III+ or tumour size ≥40 mm [g]	170032 (3 RCTs)[10,6,8,9,a]	⨁⨁◯◯ LOW[c,h]	**RR 0.62** (0.48 to 0.80)	Low	
				170 per	**65 fewer per 100.000**

				100.000[f]	(88 fewer to 34 fewer)
Other cause mortality follow up: mean 9.6 years	119083 (3 RCTs)[11,9,a]	⊕⊕◯◯ LOW[c,j]	**RR 0.99** (0.95 to 1.04)	Low	
				6.600 per 100.000[f]	**66 fewer per 100.000** (330 fewer to 264 more)
Overdiagnosis (population perspective)	64117 (2 RCTs)[12,9,a]	⊕⊕⊕◯ MODERATE[c]	-	10.1% (95% CI 8.6%-11.6%) [k]	
Overdiagnosis (patient perspective)	64117 (2 RCTs)[12,9,a]	⊕⊕⊕◯ MODERATE[c]	-	17.3% (95%CI 14.7%-20.0%) [l]	
Rate of mastectomies[g]	249550 (5 RCTs)[13,14,15,16,3,a]	⊕⊕◯◯ LOW[c,m]	**RR 1.20** (1.11 to 1.30)[n]	Low	
				900 per 100.000[f]	**180 more per 100.000** (99 more to 270 more)
Provision of chemotheraphy[g]	99454 (2 RCTs)[15,16,3,a]	⊕◯◯◯ VERY LOW[c,j,o,p]	**RR 0.86** (0.52 to 1.41)[n]	Low	
				400 per 100.000[f]	**56 fewer per 100.000** (192 fewer to 164 more)
Quality of life (inferred from psychological effects)[g]	(54 observational studies)[17]	⊕⊕◯◯ 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	(24 observational	⊕⊕◯◯	-	Two systematic reviews. One review included 17 studies and	

	effects (psychological distress)[g]	studies)[18,19]	LOW		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)[g]	(4 observational studies)[20]	⊕⊕◯◯ LOW	-	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. Bjurstam NG, Björneld LM,Duffy SW. Updated results of the Gothenburg Trial of Mammographic Screening. Cancer; 2016.
2. S, Shapiro. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan.. J Natl Cancer Inst Monogr; 1997.
3. 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.
4. 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.
5. Miller AB, Baines CJ,To T,Wall C.. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years.. CMAJ.; 1992.
6. 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.
7. Bjurstam N1, Björneld L,Warwick J,Sala E,Duffy SW,Nyström L,et al. The Gothenburg Breast Screening Trial.. Cancer; 2003.
8. 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.
9. Miller AB, To T,Baines CJ,Wall C.. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years.. J Natl Cancer Inst.; 2000.
10. 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.
11. 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 .
12. Zackrisson S, Andersson I, Janzon L,Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ; 2006.
13. J, Frisell. Mammographic screening for breast cancer [thesis]. Stockholm: Södersjukhuset; 1989.
14. AB, Miller. The costs and benefits of breast cancer screening.. Am J Prev Med; 1993.
15. 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.
16. 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.
17. Brett J, Bankhead C,Henderson B,Watson E,Austoker J.. The psychological impact of mammographic screening. A systematic review.. Psychooncology; 2005.
18. 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.
19. Salz T, Richman AR,Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes.. Psychooncology.; 2010.
20. 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. The GDG felt that baseline risks higher than 0.6% should be considered to evaluate absolute effects (Breast Cancer Screening, IARC Handbook of Cancer Prevention Volume 15) c. Trials were conducted more than 20 years ago. Currently, women have higher adherence to breast cancer screening while quality control of screening and breast cancer care have improved. d. Despite concerns about indirectness from the trials, including the fact that the population age range of 40 to 74 is broader than the age range in this question, after considering evidence from contemporary non-randomised studies (Broeders 2012) the GDG decided not to downgrade the quality of evidence for indirectness. e. 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-2 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 that 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'. f. Median or mean of the control group of the included studies unless otherwise specified. g. Importance of the outcome was lowered from 'critical' to 'important' because the members felt this outcome influenced neither the direction nor the strength of the recommendation. h. Non-blinded assessment of breast cancer stage is a serious concern. GDG members decided to downgrade to 'serious' for risk of bias. i. Unexplained inconsistency with statistical heterogeneity (I^2 = 70%, P = 0.02). While one study shows clear benefit, in three studies the 95%CI does not exclude important benefit or harm. j. 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). k. Estimate from a meta-analysis of 2 trials (CNBSS-2 and Malmo I) 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). l. Estimate from a meta-analysis of 2 trials (CNBSS-2 and Malmo I) 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). m. Observational studies do not confirm these results, instead they provide opposite results. n. 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 an increased rate of chemotherapy and mastectomies in the screened group o. Unexplained inconsistency with statistical heterogeneity (I^2 = 71%, P=0.06). p. Chemotherapy protocols and indications have significantly changed (e.g. node status was not determined in earlier studies). q. Unexplained inconsistency for variability in anxiety in the group of women recalled for further testing.	

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 moderate, as this was the lowest quality (corresponding to the quality of the evidence for overdiagnosis) of the two critical outcomes—namely, breast cancer mortality and overdiagnosis.	Effects of chemotherapy and mastectomy were not considered to change the recommendation, and thus did not critically influence the overall certainty in the evidence.

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		Two GDG members disagreed.

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	(3) found that the total screening cost for more than 400 000 women aged 50 to 69 followed for 15 years was €1 126.6 million (using a 3% discount rate). The incremental cost for a breast cancer screening programme was €36.4 million compared to no screening. (4) estimated the total costs related to diagnosis, treatment and death in the absence of screening at €1 161 008 per 1 000 women aged 50 to 74 followed over their lifetime (using a 3.5% discount rate). The incremental cost for breast cancer screening was €137 057 compared to no screening. (5) found that the total screening cost was £179 million (using a 3.5% discount rate) for a cohort of 364 500 women aged 50 to 70, followed for 35 years. The incremental cost of the screening programme was £42.5 million compared to no screening. (6) showed that the total cost for biennial screening for 100 000 women aged 50 to 69 was €143.4 ×10^6 (using a 3% discount rate). Screening was estimated to cost €16.2 ×10^6 higher than no screening. The total cost for no screening was €1 239×10^6 (using a 3% discount rate) for 1 000 000 women aged 50 to 69 (7). The incremental cost for mammography screening was €391×10^6 compared to no screening. (8) reported a total cost for breast screening of 540 996 438 FRF (using a 5% discount rate) for 315 274 women aged 50 to 65 followed for 20 years. The incremental cost of screening was FRF 199 384 259 compared to no screening. (9) found that the total cost of a national programme were $USD 154 796 590 (using a 4.5% discount rate) for 159 887 women aged 50 to 69 followed for 24 years. No costs were imputed for 'no-screening programme'. (10) indicated that the incremental cost for the breast cancer screening programme was $10 956 115 (using a 3% discount rate) for 90 000 women aged 50 to 59 followed for 27 years. No costs were assumed for 'no-screening programme'. (11) showed that the total cost for screening 60 147 women aged 50 to 69 followed for 10 years was £4 928 840. The reported incremental cost for screening was £3 492 844 compared to no screening.	Although costs were considered moderate, these costs differ by country and they are influenced by the presence of opportunistic screening.

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. On the one hand, parameters used in the model of (3), (4), (6), (7), (9), (10) and (11) were based on data from a biennial screening. On the other hand, parameters used in (5) and (12) were from a triennial screening. Only the model of (8) used parameters from an annual screening. The studies of (3), (4), (5), (6), (7), (8), (10), (11) reported costs of screening, diagnosis, and treatment. In (9), the cost per screen included cost of invitation, screening and diagnostic work-up until a final benign/malign diagnosis, but not treatment. (12) did not report cost components. The formal assessment of the certainty in the evidence for cost and resources used was made following the GRADE criteria and informed in the Evidence Profile (JRC Technical Report PICO 14-15, contract FWC443094012015; available upon request).	

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	(3) found that the screening programme during the 15-year period for women aged 50 to 69 is related to an Incremental Cost-Effectiveness Ratio (ICER) below the threshold of €30 000 per Quality Adjusted Life Year (QALY). The programme proved to be cost-effective during the evaluation phase. Based on the evidence provided by (4), the screening programme for women aged 50 to 74 was associated with an ICER that was below the cost-effectiveness threshold (£20 000 / €24 000). (5) recalculated the incremental cost of screening against the change in QALYs for each of the 5000 model runs under the base case scenario. The probability that the breast screening programme was cost effective compared with no screening was 45% (2260 scenarios) at a threshold of £20 000 per QALY. In 588 (12%) model runs, the screening programme was associated with a reduction in QALYs. Furthermore, (6) selected the biennial strategies for women aged 50 to 69 as cost-effective for both effect measures, life year gained (LYG) and QALY. The findings of (12) showed that based on commonly quoted thresholds of society's 'willingness-to-pay' of $USD 50 000 per QALY, the optimal cost-effective approach in the Slovenian population would be screening women aged 40 to 80 every three years. From the data published by (10) the ICER of biennial screening was $USD 3 750 per LYG ranging from $USD 15 502 to $USD 40 308. The authors did not specify whether the screening strategy was cost-effective. However, (9) reported an ICER of $USD 3 750 per year of life saved. The cost-effectiveness analysis shows that it is possible to run a highly cost-efficient screening programme for women aged 50 to 69. (11) found that the cost per life year gained was £ 8 561 and concluded that mammography screening was cost-effective.	(5)Cost-effectiveness probably favours the intervention in different countries or settings but varies across them. Differences in the cost-effectiveness results could be explained by the differences in screening policies, settings, outcomes and types of technology used. (4) reported the ICER per LYG, (5) and (12) considered the ICER per QALY. Whereas, (6) reported the ICER per QALY, Life Expectancy (LE) and LYG. (4) assessed digital mammography while other studies assessed screen-film mammography.

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		A systematic review on this topic has not been conducted. 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 because of 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		A systematic review on this topic has not been conducted. Some countries do not have organised screening programmes in place and may not be able to implement them mainly due to lack of resources and / or infrastructure.

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

	JUDGEMENT						
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 50 to 69 with an average risk of breast cancer, the ECIBC's Guidelines Development Group (GDG) recommends mammography screening over no mammography screening, in the context of an organised screening programme (strong recommendation, moderate certainty of the evidence).

Justification

Overall justification

The strong recommendation (rather than conditional) in favour of mammography screening over no mammography screening, in the context of an organised screening programme, was a result of a balance of the health effects that favours the intervention, in the context of moderate certainty in the evidence about these effects, and the cost-effectiveness of screening probably favouring the intervention. As agreement within the GDG for the strength of this recommendation could not be reached, voting among the members without CoI resulted in the following: 15 members voted in favour of 'strong recommendation' two members voted in favour of 'conditional recommendation', and 1 member abstained.

Detailed justification

Desirable Effects

Mammography screening, compared to no screening, reduced the risk of breast cancer mortality using short case accrual (high quality evidence). The absolute effect of mammography screening (with a mean follow-up of 18 years) varied depending on the baseline risk considered. The GDG examined three different breast cancer baseline risks: - For 2.1% , the absolute effect of screening was 462 fewer breast cancer deaths per 100 000 (with a range from 168 to 714). - For 1%, the absolute effect of screening was 220 fewer breast cancer deaths per 100 000 (with a range from 80 to 340); - For 0.6%, the absolute effect of screening was 132 fewer breast cancer deaths per 100 000 (with a range from 48 to 204); Other methods for calculating breast cancer mortality may be used (more information can be found in the annex with the evidence profile). Mammography screening also reduced breast cancer mortality using longest case accrual available (167 fewer breast cancer deaths per 100 000 women over 17.3 years) (high 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 10.1% (moderate quality evidence) from a population perspective and 17.3% from the perspective of a woman invited to screening (moderate quality evidence). Women who had further testing following their routine mammogram experienced significant short-term anxiety. Estimated cumulative risk of a false-positive screening result in women aged 50 to 69 undergoing 10 biennial screening tests was 19.7%, with 2.2% of women having a needle biopsy after an 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).

Cost effectiveness
Cost-effectiveness probably favours the intervention in different countries or settings but varies across them. Differences in the cost-effectiveness results could be explained by the differences in setting, outcomes and type of technology used.

Subgroup considerations

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

Implementation considerations

Despite being a strong recommendation, women should be provided with the information regarding benefits and harms of screening.

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. Further research on age-specific effects related to benefits and harms is needed.

2. Better information/evidence about overdiagnosis is needed.

3. A better understanding of the natural history of breast cancer.

4. Stratification possibilities.

5. Use of monitoring data to assess effectiveness (see monitoring).

6. Role of breast density in stratification.

REFERENCES SUMMARY

1. Ferlay J, Steliarova-Foucher E,Lortet-Tieulent J,Rosso S,Coebergh JW,Comber H,Forman D,Bray F.. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer; 2013.
2. Tsilidis KK, Papadimitriou N,Capothanassi D,Bamia C,Benetou V,Jenab M,Freisling H,Kee F,Nelen A,O'Doherty MG,Scott A,Soerjomataram I,Tjonneland A,May AM,Ramon Quiros J,Pettersson-Kymmer U,Brenner H,Schottker B,Ordonez-Mena JM,Karina Dieffenbach A,Eriksson S,Bogeberg Mathiesen E,Njolstad I,Siganos G,Wilsgaard T,Boffetta P,Trichopoulos D,Trichopoulou A.. Burden of Cancer in a Large Consortium of Prospective Cohorts in Europe. J Natl Cancer Inst; 2016.
3. Arrospide A, Rue M,van Ravesteyn NT,Comas M,Soto-Gordoa M,Sarriugarte G,Mar J. Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis. BMC Cancer; 2016.
4. Sankatsing VD, Heijnsdijk EA,van Luijt PA,van Ravesteyn NT,Fracheboud J,de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. Int J Cancer; 2015.
5. Pharoah PD, Sewell B,Fitzsimmons D,Bennett HS,Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. BMJ; 2013.
6. Carles M, Vilaprinyo E,Cots F,Gregori A,Pla R,Román R,Sala M,Macià F,Castells X,Rue M. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer; 2011.
7. de Gelder R, Bulliard JL,de Wolf C,Fracheboud J,Draisma G,Schopper D,de Koning HJ. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer; 2009.
8. Arveux P, Wait S,Schaffer P. Building a model to determine the cost-effectiveness of breast cancer screening in France. Eur J Cancer Care (Engl); 2003.
9. Wang H, Kåresen R,Hervik A,Thoresen SO. Mammography screening in Norway: results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes Control; 2001.
10. Leivo T, Sintonen H,Tuominen R,Hakama M,Pukkala E,Heinonen OP. The cost-effectiveness of nationwide breast carcinoma screening in Finland, 1987-1992. Cancer; 1999.
11. Norum, J. Breast cancer screening by mammography in Norway. Is it cost-effective?. Ann Oncol; 1999.
12. Rojnik K, Naversnik K,Mateović-Rojnik T,Primiczakelj M. Probabilistic cost-effectiveness modeling of different breast cancer screening policies in Slovenia. Value Health; 2008.