Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Atakpa EC, Thorat MA, Cuzick J, Brentnall AR

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

i

[Prognosis Review]

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review

Emma C Atakpa[1], Mangesh A Thorat[1,2,3], Jack Cuzick[1], Adam R Brentnall[1]

[1]Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [2]School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, London, UK. [3]Breast Services, Guy's Hospital, Guy's and St Thomas' NHS Foundation Trust, London, UK

Contact address: Adam R Brentnall, a.brentnall@qmul.ac.uk.

A B S T R A C T

Background

Endocrine therapy is effective at preventing or treating breast cancer. Some forms of endocrine therapy have been shown to reduce mammographic density. Reduced mammographic density for women receiving endocrine therapy could be used to estimate the chance of breast cancer returning or developing breast cancer in the first instance (a prognostic biomarker). In addition, changes in mammographic density might be able to predict how well a woman responds to endocrine therapy (a predictive biomarker). The role of breast density as a prognostic or predictive biomarker could help improve the management of breast cancer.

Objectives

To assess the evidence that a reduction in mammographic density following endocrine therapy for breast cancer prevention in women without previous breast cancer, or for treatment in women with early-stage hormone receptor-positive breast cancer, is a prognostic or predictive biomarker.

Search methods

We searched the Cochrane Breast Cancer Group Specialised Register, CENTRAL, MEDLINE, Embase, and two trials registers on 3 August 2020 along with reference checking, bibliographic searching, and contact with study authors to obtain further data.

Selection criteria

We included randomised, cohort and case-control studies of adult women with or without breast cancer receiving endocrine therapy. Endocrine therapy agents included were selective oestrogen receptor modulators and aromatase inhibitors. We required breast density before start of endocrine therapy and at follow-up. We included studies published in English.

Data collection and analysis

We used standard methodological procedures expected by Cochrane. Two review authors independently extracted data and assessed risk of bias using adapted Quality in Prognostic Studies (QUIPS) and Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) tools. We used the GRADE approach to evaluate the certainty of the evidence. We did not perform a quantitative meta-analysis due to substantial heterogeneity across studies.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1

Main results

Eight studies met our inclusion criteria, of which seven provided data on outcomes listed in the protocol (5786 women). There was substantial heterogeneity across studies in design, sample size (349 to 1066 women), participant characteristics, follow-up (5 to 14 years), and endocrine therapy agent. There were five breast density measures and six density change definitions. All studies had at least one domain as at moderate or high risk of bias. Common concerns were whether the study sample reflected the review target population, and likely post hoc definitions of breast density change.

Most studies on prognosis for women receiving endocrine therapy reported a reduced risk associated with breast density reduction. Across endpoints, settings, and agents, risk ratio point estimates (most likely value) were between 0.1 and 1.5, but with substantial uncertainty. There was greatest consistency in the direction and magnitude of the effect for tamoxifen (across endpoints and settings, risk ratio point estimates were between 0.3 and 0.7). The findings are summarised as follows.

Prognostic biomarker findings:

Treatment

Breast cancer mortality

Two studies of 823 women on tamoxifen (172 breast cancer deaths) reported risk ratio point estimates of ~0.4 and ~0.5 associated with a density reduction. The certainty of the evidence was low.

Recurrence

Two studies of 1956 women on tamoxifen reported risk ratio point estimates of ~0.4 and ~0.7 associated with a density reduction. There was risk of bias in methodology for design and analysis of the studies and considerable uncertainty over the size of the effect.

One study of 175 women receiving an aromatase inhibitor reported a risk ratio point estimate of ~0.1 associated with a density reduction. There was considerable uncertainty about the effect size and a moderate or high risk of bias in all domains.

One study of 284 women receiving exemestane or tamoxifen as part of a randomised controlled trial reported risk ratio point estimates of ~1.5 (loco-regional recurrence) and ~1.3 (distance recurrence) associated with a density reduction. There was risk of bias in reporting and study confounding, and uncertainty over the size of the effects.

The certainty of the evidence for all recurrence endpoints was very low.

Incidence of a secondary primary breast cancer

Two studies of 451 women on exemestane, tamoxifen, or unknown endocrine therapy reported risk ratio point estimates of ~0.5 and ~0.6 associated with a density reduction. There was risk of bias in reporting and study confounding, and uncertainty over the effect size. The certainty of the evidence was very low.

We were unable to find data regarding the remaining nine outcomes prespecified in the review protocol.

Prevention

Incidence of invasive breast cancer and ductal carcinoma in situ (DCIS)

One study of 507 women without breast cancer who were receiving preventive tamoxifen as part of a randomised controlled trial (51 subsequent breast cancers) reported a risk ratio point estimate of ~0.3 associated with a density reduction. The certainty of the evidence was low.

Predictive biomarker findings:

One study of a subset of 1065 women from a randomised controlled trial assessed how much the effect of endocrine therapy could be explained by breast density declines in those receiving endocrine therapy. This study evaluated the prevention of invasive breast cancer and DCIS. We found some evidence to support the hypothesis, with a risk ratio interaction point estimate ~0.5. However, the 95% confidence interval included unity, and data were based on 51 women with subsequent breast cancer in the tamoxifen group. The certainty of the evidence was low.

Authors' conclusions

There is low-/very low-certainty evidence to support the hypothesis that breast density change following endocrine therapy is a prognostic biomarker for treatment or prevention. Studies suggested a potentially large effect size with tamoxifen, but the evidence was limited. There was less evidence that breast density change following tamoxifen preventive therapy is a predictive biomarker than prognostic biomarker. Evidence for breast density change as a prognostic treatment biomarker was stronger for tamoxifen than aromatase inhibitors. There were no studies reporting mammographic density change following endocrine therapy as a predictive biomarker in the treatment setting, nor

aromatase inhibitor therapy as a prognostic or predictive biomarker in the preventive setting. Further research is warranted to assess mammographic density as a biomarker for all classes of endocrine therapy and review endpoints.

PLAIN LANGUAGE SUMMARY

Reduced breast density following endocrine therapy as an indicator of breast cancer risk

What is the issue?

Breast cancer is a common cancer and cause of death in women worldwide. Treatment options for breast cancer include endocrine therapy. Endocrine therapy can also be used to prevent breast cancer for women who have not been diagnosed with breast cancer. It would help doctors and their patients to understand whether some patients are likely to have greater benefit from endocrine therapy than others. The structure of the breast is likely to change following endocrine therapy. These structural changes are seen when women have a mammogram (breast x-ray). They appear as a decrease in the area of white tissue (breast density) on the mammogram. We wanted to find out whether reductions in breast density after endocrine therapy can help to determine how well endocrine therapy works.

Review question

We searched for previously published studies. We assessed whether a reduction in breast density after receiving endocrine therapy was associated with better outcomes. For women without breast cancer, this focused on whether those with decreased breast density were less likely to develop breast cancer. For women with breast cancer, this included whether those with greater decreases in breast density were less likely to die from breast cancer.

Study characteristics

We performed the search on 3 August 2020. We included studies of adult women with breast cancer if the women's breast cancer had been diagnosed at an early stage and could be treated with endocrine therapy (hormone receptor-positive breast cancer). We included drugs often used in practice (tamoxifen and aromatase inhibitors). We found a wide variety of studies. The studies varied in terms of how they had been planned and the characteristics of the women included in the studies, as well as in how breast density change was measured.

Key results

Most studies reported a reduced risk of breast cancer after endocrine therapy for women who had a breast density reduction compared with women who did not have a reduction. There was slightly stronger evidence for the drug tamoxifen.

• Two studies reported on breast density reduction following tamoxifen and risk of breast cancer death. The findings were based on 172 women who died from breast cancer. Overall, the certainty of the evidence was low.

• Two studies considered if breast cancer returned after treatment with tamoxifen. There were concerns about the study methods and certainty of findings in these two studies. Overall, the certainty of the evidence was very low.

• One study considered treatment with an aromatase inhibitor and the chance of breast cancer returning. There was considerable uncertainty about the effect size because there were only 175 women in the study. The certainty of the evidence was very low due to potential risk of bias in the study.

• One study considered if breast cancer returned locally or at a distance from the original tumour. There was risk of bias in reporting and uncertainty about the sizes of the effect. The certainty of the evidence for both outcomes was very low.

• Two studies looked at the chance of women with breast cancer being diagnosed later with a new breast cancer, such as in the opposite breast. There was risk of bias in reporting and uncertainty about the size of the effect. The certainty of the evidence was very low.

• One study considered women who had not previously had breast cancer and who received tamoxifen. Results were based on 51 women who developed breast cancer. Overall, the certainty of the evidence was low.

• One study considered whether the beneficial effect of tamoxifen could be explained by a decrease in breast density. There was some evidence to support this, but there was uncertainty about the strength of the effect. The results were based on 51 women who developed breast cancer after receiving tamoxifen. The certainty of the evidence was low.

Overall, we found some evidence that breast density change following tamoxifen therapy is useful information to help determine how well the drug will work in future. However, there is much uncertainty about the strength of this effect. This was due to small numbers of women in the studies, relatively few studies for each outcome, and limitations in many of the studies such as how breast density change was measured. More research is needed to help assess these issues.

Quality of the evidence

Overall, we assessed the certainty of the available evidence as low or very low.

SUMMARY OF FINDINGS

Summary of findings 1. Prognostic biomarker - treatment studies

Women receiving endocrine therapy: mammographic density reduction versus no mammographic density reduction

Patient or population: women receiving endocrine therapy (selective oestrogen receptor modulator or aromatase inhibitor)
Setting: treatment of early-stage breast cancer
Intervention: mammographic density reduction
Comparison: no mammographic density reduction

Outcomes	Impact	№ of participants (studies)	Certainty of the evidence (GRADE)
Breast cancer mortality - tamoxifen	OR 0.44 (95% CI 0.22 to 0.88) for per cent density reduction > 8.7% compared with < 0.5% (Cumulus density from mammograms 3 to 26 months apart; Nyante 2015). HR 0.50 (95% CI 0.27 to 0.93) for relative reduction in dense area > 20% compared with ≤ 9% increase to ≤ 10% reduction (automated machine-learning measure on mammograms 6 to 36 months apart; Li 2013).	97 cases/252 controls (1 study) 217 exposed (26 events)/257 unexposed (49 events) (1 study)	⊕⊕⊖⊖ LOW [1]
Recurrence - tamoxifen	HR 0.66 (95% CI 0.40 to 1.09) for per cent density reduction ≥ 5% compared with < 5% (Cumulus density from mammograms 8 to 20 months apart; Kim 2012). HR 0.35 (95% CI 0.17 to 0.68) for reduction in BI-RADS density categories compared with no reduction (BI-RADS density on mammograms approximately 10 to 34 months apart; Ko 2013).	1956 (2 studies)	⊕⊖⊖⊖ VERY LOW [1 2 3]
Recurrence - AIs	HR 0.14 (95% CI 0.02 to 1.11) for per cent density reduction ≥ 5% compared with < 5% (Cumulus density from mammograms 8 to 20 months apart; Kim 2012).	175 (1 study)	⊕⊖⊖⊖ VERY LOW [1 4 5]
Recurrence (loco-regional) - endocrine therapies not separated (exemestane (AI)/tamoxifen (SERM))	HR 1.48 (95% CI 0.38 to 5.81) per Boyd category reduction (visual assessment of Boyd categories from mammograms 18 to 30 months apart; van Nes 2015).	284 (1 study)	⊕⊖⊖⊖ VERY LOW [6 7]
Recurrence (distance) - endocrine therapies not separated (exemestane (AI)/tamoxifen (SERM))	HR 1.32 (95% CI 0.64 to 2.71) per Boyd category reduction (visual assessment of Boyd categories from mammograms 18 to 30 months apart; van Nes 2015).	284 (1 study)	⊕⊖⊖⊖ VERY LOW [6 7]
Incidence of a secondary primary breast cancer (e.g. in the contralateral breast) - endocrine therapies not separated (exemes-	OR 0.52 (95% CI 0.18 to 1.51) for per cent density reduction ≥ 10% compared with < 10% increase to < 10% reduction (automated machine-learning measure on mammograms 12 to 60 months apart; unknown endocrine therapy; Sandberg 2013). HR 0.58 (95% CI 0.08 to 4.44) per Boyd category reduction (visual assessment of Boyd categories from mammograms 18 to 30 months apart; exemestane or tamoxifen; van Nes 2015).	87 cases/87 controls (1 study) 277 (1 study)	⊕⊖⊖⊖ VERY LOW [8 9]

tane (AI)/tamoxifen
(SERM)/unknown)

AI: aromatase inhibitor; **BI-RADS**: Breast Imaging Reporting and Data System; **CI**: confidence interval; **DCIS**: ductal carcinoma in situ; **HR:** hazard ratio; **OR:** odds ratio; **SERM**: selective oestrogen receptor modulator

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Certainty of evidence was initially graded low due to the observational, non-randomised design of the contributing studies. No downgrading was applied.

[2]We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - risk of bias): moderate risk of bias due to study participation, attrition, prognostic factor measurement, and outcome measurement in both studies; it is unclear if there were adjustments made in one study and no adjustment for important confounding factors such as chemotherapy made in another study; high risk of bias in statistical analysis for both studies due to potential difference in immortal time included in analysis (greater for those with larger density reductions).

[3]We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - imprecision): confidence interval includes the null effect for the first study; in total, the evidence includes no more than 147 events (first study did not report the number of events in the tamoxifen subgroup (80 events in the study overall); 67 events in the second study).

[4]We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - risk of bias): moderate risk of bias due to study participation, attrition, prognostic factor measurement, and outcome measurement; unclear if adjustments were made, and unclear if some women were treated with tamoxifen between baseline and follow-up mammogram; high risk of bias in statistical analysis due to potential difference in immortal time included in analysis (greater for those with larger density reductions).

[5]We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - imprecision): confidence interval includes the null effect; no more than 80 events (study does not provide the number of events in the AI subgroup, but there were 80 events in the study overall).

[6]Certainty of evidence was initially graded moderate because data for these outcomes were from a randomised controlled trial. We downgraded the certainty of the evidence from moderate to very low (GRADE downgrading factor - risk of bias): the analysis of interest was a secondary objective and included only a subgroup of participants (women still at risk after two years of therapy), therefore data on women included in the subgroup analysis are not reported; results by individual endocrine therapy are not reported; there was no adjustment for important confounding factors such as age, body mass index, baseline density, time between mammograms, and chemotherapy.

[7]Certainty of evidence was initially graded moderate because data for these outcomes were from a randomised controlled trial. We downgraded the certainty of the evidence from moderate to very low (GRADE downgrading factor - imprecision): confidence interval includes the null effect; no more than 284 events (study does not provide the number of events, but there were 284 women overall in the subgroup of women still at risk after 2 years of therapy).

[8]Certainty of evidence was initially graded low due to the observational, non-randomised design of one of the contributing studies. We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - risk of bias): in both studies, the analysis of interest was a secondary objective and included only a subgroup of participants (women on endocrine therapy and women still at risk after two years of therapy), therefore data on women included in the subgroup analyses are not reported; results by individual endocrine therapy are not reported (endocrine therapy was unknown in one study, and women were on either exemestane or tamoxifen between mammograms in the other study); one study matched for some confounding factors, but it was unclear if further adjustments were made, and no adjustments were made for important confounding factors such as age, body mass index, baseline density, time between mammograms, and chemotherapy in the other study.

[9]Certainty of evidence was initially graded low due to the observational, non-randomised design of one of the contributing studies. We downgraded the certainty of the evidence from low to very low (GRADE downgrading factor - imprecision): confidence intervals include the null effect; no more than 364 events (one study does not provide the number of events by density change category, but there were 87 cases overall on endocrine therapy; the other study does not provide the number of events, but there were 277 women overall in the subgroup of women still at risk after 2 years of therapy).

Summary of findings 2.   Prognostic biomarker - prevention studies

Women receiving endocrine therapy: mammographic density reduction versus no mammographic density reduction

Patient or population: women receiving endocrine therapy (selective oestrogen receptor modulator or aromatase inhibitor)
Setting: prevention of breast cancer
Intervention: mammographic density reduction
Comparison: no mammographic density reduction

Outcomes	Impact	№ of participants (studies)	Certainty of the evidence (GRADE)
Incidence of invasive breast cancer and DCIS - tamoxifen	OR 0.32 (95% CI 0.14 to 0.72) for per cent density reduction ≥ 10% compared with 0% density change (visual assessment in 5% bands from mammograms 12 to 18 months apart; Cuzick 2011a)	51 cases/456 controls (1 study)	⊕⊕⊖⊖ LOW [1]

CI: confidence interval; **DCIS:** ductal carcinoma in situ; **OR:** odds ratio

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Certainty of evidence was initially graded moderate because data for this outcome were from a randomised controlled trial. We downgraded the certainty of the evidence from moderate to low (GRADE downgrading factor - risk of bias): cutpoint for the mammographic density biomarker was not determined a priori, but was based on the observed study data.

Summary of findings 3. Predictive biomarker

Women receiving endocrine therapy or not receiving endocrine therapy (control group): effect of mammographic density reduction in endocrine therapy group versus effect of mammographic density reduction in control group

Patient or population: women receiving endocrine therapy (selective oestrogen receptor modulator or aromatase inhibitor) or not receiving endocrine therapy (control group)
Setting: prevention of breast cancer or treatment of early-stage breast cancer
Intervention: effect of mammographic density reduction in endocrine therapy group
Comparison: effect of mammographic density reduction in control group

Outcomes	Impact	№ of participants (studies)	Certainty of the evidence (GRADE)
Prevention: incidence of invasive breast cancer and DCIS - tamoxifen	OR 0.53 (95% CI 0.21 to 1.32) for an interaction between per cent density reduction (≥ 10% or < 10%) and prophylactic tamoxifen (visual assessment in 5% intervals on mammograms 12 to 18 months apart; Cuzick 2011a)	123 cases/942 controls (1 observational study)	⊕⊕⊖⊖ LOW [1]

CI: confidence interval; **DCIS:** ductal carcinoma in situ; **OR:** odds ratio

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

6

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

---

[1]Certainty of evidence was initially graded moderate because data for this outcome were from a randomised controlled trial. We downgraded the certainty of the evidence from moderate to low (GRADE downgrading factor - imprecision): confidence interval is wide and includes the null effect; power to detect an interaction is limited by the sample size and number of cases.

BACKGROUND

Description of the condition and intervention

Breast cancer is the most common cancer in women worldwide, the second most frequent cause of cancer death in women from high-income regions, and the most common cause of death in low-income regions (Ferlay 2013). Two types of drugs have shown efficacy for both prevention and treatment of certain subtypes of the disease. The first are selective oestrogen receptor modulators (SERMS), which prevent breast cancer (Cuzick 2013; Cuzick 2015), and are also used in adjuvant settings to reduce the chance that breast cancer will reoccur when it has been diagnosed at an early stage (Davies 2011; EBCTCG 1998). The second are aromatase inhibitors (AIs), which are suitable for postmenopausal women only, and confer greater average reductions in the risk of breast cancer, Cuzick 2014; Goss 2011; Visvanathan 2013, and recurrence than SERMs (EBCTCG 2015).

Description of the biomarker

The breast is made up of glandular and supportive tissue. Glandular tissue is the network that produces and transports milk to the nipple; the supportive tissue is largely fat but also contains fibrocollagenous tissue known as glandular stroma. On a mammogram (breast x-ray), glandular tissue and glandular stroma appear as a white area which is referred to as mammographic density (Assi 2011). Breast density is a strong risk factor for breast cancer, and women with mostly dense breasts have approximately four times the risk of breast cancer than women of the same age and weight with mostly fatty breasts (Huo 2014; McCormack 2006). Mammographic density is also associated with classical reproductive risk factors, and is lower in women who have had children and who have breastfed (Boyd 1998).

How the biomarker might be related to treatment response

Hormonal treatment can change a woman's mammographic density. Density increases during the use of hormone replacement therapy (HRT), and HRT is also a risk factor for breast cancer (McTiernan 2005; Rutter 2001). After cessation of HRT, mammographic density may decrease in as little as four weeks (Harvey 1997); however, excess risk may persist for more than 10 years after cessation, the magnitude of which depends on the duration of previous use, with little excess risk observed following less than 1 year of HRT use (CGHFBC 2019). Breast density may also decrease during SERM therapy above that expected due to age (Cuzick 2004), whilst the evidence for AIs is less clear (Engmann 2017; Vachon 2013b).

The association between hormonal treatment and density change is well documented, and there is also direct evidence that the increased risk from combination HRT is mediated by mammographic density (Boyd 2006; Byrne 2017; Martin 2009). Findings for prevention, Cuzick 2011a, and treatment (including Kim 2012; Ko 2013; Li 2013; Nyante 2015), also suggest that change in breast density is an appropriate biomarker for response to SERMs. A working hypothesis is therefore that mammographic density reductions in women receiving endocrine therapy for treatment or prevention might indicate who is responding to therapy, making it a reliable surrogate outcome. The precise mechanism is still unclear and is an area of active research, but one theory is that decreases in density arise when a woman is able to metabolise the drug effectively (Jordan 2007).

Why it is important to do this review

The first aim of this review was to assess the evidence that change in mammographic density is a prognostic biomarker (Altman 2001). We define the term 'prognostic biomarker' to be a measure that is associated with a clinical outcome of interest in a defined group of patients. This terminology is standard when the group of patients has a health condition such as breast cancer, but is perhaps less frequently used for risk factors in healthy patients when the clinical outcome is breast cancer.

Several prognostic factors for women diagnosed with breast cancer have been identified. These include classical factors such as tumour size, grade and lymph node involvement, and biomarkers including Ki67 and commercial genetic signatures such as Oncotype DX (Cuzick 2011b; Harris 2007). Prognostic factors for healthy women without breast cancer (or risk factors) include age, a family history of the disease, and hormonal and reproductive factors including weight and age at first child (Tyrer 2004). Quantifying the effect of potential prognostic factors on outcomes is important for many reasons. It may be used to help guide clinical decision making and improve the understanding of disease, the design and analysis of trials, and risk assessment (Riley 2013).

The second aim of this review was to assess the evidence that change in mammographic density is a predictive biomarker, which is taken to be a measure that is differentially associated with response to treatment (Hingorani 2013). Some, but not all, prognostic biomarkers are predictive biomarkers. Two examples for women with breast cancer are human epidermal growth factor receptor 2 (HER-2) and oestrogen receptor (ER) status. HER-2 was identified as a prognostic factor for breast cancer and provided a target for a treatment (trastuzumab) that was subsequently shown to be effective for women with HER-2 breast cancer. ER status is a prognostic biomarker and a predictive biomarker for SERM and AI treatments, that is they have been shown to improve clinical outcomes only in ER-positive patients.

There is currently no systematic review focusing on the evidence that mammographic density reductions in women receiving endocrine therapy are prognostic or predictive biomarkers. However, some other reviews on the topic have been published, including the Shawky 2017 study. Shawky 2017 reported seven studies of density change as a prognostic factor for women receiving a SERM or AI, but no data from a randomised trial or otherwise to evaluate change in mammographic density after initiation of adjuvant tamoxifen treatment as a predictive biomarker. There has been one study to evaluate density change as a prognostic and predictive biomarker for prevention, which was a case-control study from within a randomised trial.

It was important to undertake this review because findings are likely to be important to clinicians and their patients undergoing or considering endocrine therapy, such as by helping to define risk groups and better predict outcomes; regulators and ethics boards considering trials of products that use mammographic density reductions as an endpoint; and those with an interest in the mechanisms by which endocrine therapy improves clinical outcomes. Additionally, as discussed in the Mullooly 2016 study, had the randomised trials of SERMS and AIs included density

change as a potential prognostic or predictive biomarker, then different conclusions might have been reached regarding their effectiveness; it is possible that women with density reductions from a SERM might receive greater benefits from this treatment than from an AI. Another possibility is that women who see density increases following a short-term decrease might show resistance to the treatment.

OBJECTIVES

Endocrine therapy for breast cancer prevention has been shown to reduce risk, and for treatment of early-stage ER-positive breast cancer to reduce breast cancer mortality. The objective of this review was to synthesise the available evidence on whether mammographic density reduction in these settings is (i) a prognostic biomarker or (ii) a predictive biomarker, as defined above. We planned to explore sources of heterogeneity to identify the impact of differences in participants, measures of mammographic density, follow-up length, and study design. Within the prognostic and predictive biomarker reviews, our analysis would consider prevention and treatment populations separately, and within these, SERMs and AIs separately.

METHODS

This review was written according to PRISMA guidelines (Liberati 2009), whilst supplemented as necessary for a predictive and prognostic biomarker review, and followed the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) guidelines (Altman 2012; McShane 2005). We conducted a literature-based analysis to identify relevant studies and then considered the application of meta-analytic methods.

Criteria for considering studies for this review

We included studies with the following designs, participants, interventions, biomarkers, and outcomes.

Types of studies

We included the same study designs for both prognostic and predictive reviews. We allowed for the inclusion of randomised and non-randomised observational studies (prospective and retrospective cohort and case-control studies). In the protocol for this review, we allowed for separate treatment of exploratory biomarker studies in the analysis, where density is one of several biomarkers considered simultaneously, but this was not identified in any of the included studies.

Targeted population

The same type of participant was included in the target population for prognostic and predictive biomarker reviews. We included all adult women aged 18 years or more, with or without breast cancer (denoted respectively as treatment, prevention), based on the following criteria.

Treatment: women with early-stage hormone receptor- (oestrogen (ER) or progesterone (PgR)) positive breast cancer. We defined this as women who have had histologically proven operable invasive hormone receptor-positive breast cancer or ductal carcinoma in situ (DCIS), and were candidates to receive endocrine adjuvant therapy; there was no clinical evidence of metastatic disease. Women were ineligible if breast density measurements were not

possible on a contralateral breast or if they had bilateral breast cancer.

Prevention: women who have not previously been diagnosed with invasive breast cancer or DCIS. There were no exclusions for level of increased risk due to genetic factors (including BRCA1 DNA repair associated/BRCA2 DNA repair associated (*BRCA1/BRCA2*) gene mutations or a family history of the disease, or both) or otherwise assessed by an absolute or relative risk prediction model. The target population would exclude women with breast implants or those who had undergone risk-reducing mastectomies because in such cases accurate breast density estimation is not possible.

Women must have been at risk for at least the length of time between baseline and follow-up mammogram. We included women who might have changed treatment or discontinued treatment throughout follow-up, but excluded women who changed treatment between the mammograms for density change (we did not exclude those who discontinued). We excluded women who received another SERM or AI before treatment.

For AI comparisons, the target population requires women to have been postmenopausal at the start of treatment, whilst for SERM comparisons they may have been pre- or postmenopausal. The target population of postmenopausal women included women who had undergone a bilateral oophorectomy; were aged more than 60 years; or were aged 40 to 59 years with an intact uterus and amenorrhoeic for at least 12 months. We excluded from our target population women who were rendered temporarily postmenopausal through medical interventions (e.g. gonadotropin-releasing hormone (GnRH) analogues).

We planned to include studies that included subsets of relevant participants in the main analysis, provided results were provided for the subset that included the relevant participants.

Types of interventions

Interventions

We defined the same types of intervention for the prognostic and predictive biomarker reviews.

We included women receiving SERMs at the following minimum doses (Komm 2014): tamoxifen, 20 mg daily; raloxifene, 60 mg daily; lasofoxifene, 0.25 mg daily; arzoxifene, 20 mg daily; droloxifene, 40 mg daily; bazedoxifene, 20 mg daily; and fulvestrant, 250 mg monthly.

We included women receiving AIs at the following minimum doses: anastrozole, 1 mg daily; letrozole, 2.5 mg daily; and exemestane, 25 mg daily. All treatments were oral, except fulvestrant (intramuscular). Women were to have received treatment for at least the length of time between baseline and follow-up mammogram (i.e. at least one year). We planned to include studies of women receiving doses lower than these doses for a secondary dose-response analysis, but would exclude them from the main analysis; no such studies were identified. We planned to include studies that were a mix of women including SERMs and AIs in the primary analysis if we were able to separate the results; otherwise, we planned to only include them in secondary analyses.

Co-interventions

We permitted the same types of co-intervention for the prognostic and predictive biomarker reviews.

For treatment, women were ineligible if they had not completed primary locoregional (surgery or radiotherapy, or both) treatment and systemic (chemotherapy or targeted therapy) treatment (where indicated) with curative intent (either in neoadjuvant or in adjuvant settings). Women were ineligible if there was a gap of more than eight weeks between different treatment interventions, for example between surgery and start of radiotherapy in those not receiving adjuvant chemotherapy. Women were also ineligible if they had received endocrine therapy for breast cancer prevention before diagnosis of breast cancer, or if endocrine treatment was started before surgery and received for more than 28 days.

We planned to include studies if some women use or used (up to two years before baseline) HRT (prevention and treatment), but planned to note this, including in the risk of bias assessment. We permitted other co-interventions, including exercise and diet advice, but planned to identify them where possible, including in the risk of bias assessment.

Comparators

The main difference between the prognostic and predictive biomarker review was the comparator.

Prognostic biomarker review

The comparison was within each intervention group (SERM or AI), where the outcome was related to the change in density over the period. This was done to help assess whether the biomarker was associated with the outcome in those receiving SERM or AI interventions, that is a prognostic biomarker.

Predictive biomarker review

The predictive biomarker review made a comparison between the intervention group and a control group from the same study. The within-study comparator group was a corresponding randomised placebo group, or a non-randomised control group of women not receiving endocrine therapy.

Biomarker

We used the same definition of biomarker for the prognostic and predictive reviews.

A measure of mammographic density was required at baseline (start of endocrine therapy or study entry in those from the control group) and follow-up. We included studies with baseline mammograms obtained before or after diagnosis (but no more than two years before diagnosis) and before the start of therapy (or study entry), and a follow-up mammogram performed 90 days to three years after the start of therapy (or study entry), with the density closest to one year from the start of endocrine therapy (or study entry), if there was a choice. We planned to record the range and average time between baseline mammogram and diagnosis, between diagnosis and start of endocrine therapy (or study entry), and between start of endocrine therapy (or study entry) and the follow-up mammogram.

We included any density method that has been shown in more than one study, outside of the review studies, to have a relationship with breast cancer risk. We planned this to include, but not be limited to, the following percentage methods: (i) visual assessment by expert in 5% bands; (ii) visual assessment by expert in percentage bands (Boyd categories); (iii) visual assessment by expert as continuous percentage (%); (iv) semi-automated thresholding such as using Cumulus software by expert (or trained) reader (Byng 1994); (v) fully automated (based on area of density); and (vi) fully automated volumetric percentage (e.g. Volpara, Highnam 2010). We additionally planned to consider the following categorical measures: (i) Breast Imaging Reporting and Data System (BI-RADS) density (D'Orsi 2013); (ii) Wolfe grade (Wolfe 1976); and (iii) Tabar grade (Gram 1997). We also considered absolute dense area or volume from: (i) semi-automated methods (including Cumulus); (ii) automated area-based methods; and (iii) fully automated volumetric methods.

We planned to consider information on reliability of density measures, including correlation between repeated measures from repeat mammograms, intraclass correlation coefficients, and Bland-Altman limits of agreement (Bland 1999), whether different interpreters of density were used; the same interpreter assessed density for each woman; the reader was blinded to case status; the reader was blinded to treatment allocation; randomisation was per mammogram (mammograms read independently) or per woman (mammograms for each woman read with the knowledge of her other mammograms); and the order of per-woman mammograms was sequential or random and assessed one at a time or simultaneously. We planned to use these for a qualitative assessment of potential bias due to measurement of the biomarker.

We did not include women or studies that had different definitions or measures of mammographic density between the time points used to assess change.

Types of outcomes to be predicted

We used the same outcome measures for the prognostic and predictive reviews.

Primary outcomes

Potential benefits from treatment

- Treatment: breast cancer mortality (time to death caused by breast cancer)
- Prevention: incidence of invasive breast cancer and DCIS

Potential harms from treatment

- Treatment and prevention: rate of all serious adverse events. These included serious side effects noted for tamoxifen (cataracts, pulmonary embolism or deep vein thrombosis, and endometrial cancer) and anastrozole (osteoporosis and bone fractures).

Secondary outcomes

Potential benefits from treatment

- Treatment: recurrence
- Treatment: incidence of a secondary primary breast cancer (e.g. in the contralateral breast)
- Treatment: any recurrence or any death (disease-free survival)
- Treatment: distant metastases

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

10

- Treatment: death from all causes (all-cause mortality)
- Treatment: recurrence of invasive cancer only
- Treatment: recurrence of DCIS cancer only
- Prevention: incidence of invasive cancer only
- Prevention: incidence of DCIS cancer only

Potential harms from treatment

- Treatment and prevention: troublesome but not serious side effects observed for SERMs and AIs, including vasomotor symptoms and joint or muscle pain.

Summary of findings table for assessing the certainty of the evidence

We produced different summary of findings tables for the prognostic and predictive biomarker reviews, but which were based on the same outcomes. We applied the methods of the GRADE approach (Schünemann 2011), using GRADEpro GDT software (GRADEpro GDT). We considered the following seven main outcomes.

- Treatment: breast cancer mortality (time to death caused by breast cancer).
- Prevention: incidence of invasive and DCIS.
- Treatment and prevention: the rate of all serious adverse events. These included serious side effects noted for tamoxifen (cataracts, pulmonary embolism or deep vein thrombosis, and endometrial cancer) and anastrozole (osteoporosis and bone fractures).
- Treatment: recurrence.
- Treatment: any recurrence or any death (disease-free survival).
- Treatment: death from all causes (all-cause mortality).
- Treatment and prevention: troublesome but not serious side effects observed for SERMs and AIs, including vasomotor symptoms and joint or muscle pain.

Search methods for identification of studies

Electronic searches

We searched the following databases on 3 August 2020.

- The Cochrane Breast Cancer Group Specialised Register. Details of the search strategies used by the Group for the identification of studies and the procedure used to code references are outlined in the Group's website (breastcancer.cochrane.org/specialised-register). We extracted and considered for inclusion in the review trials with the following key words: Tamoxifen, Raloxifene, Lasofoxifene, Arzoxifene, Droloxifene, Bazedoxifene, Fulvestrant, Anastrozole, Letrozole, Exemestane, selective estrogen receptor modulator, aromatase inhibitor.
- The Cochrane Central Register of Controlled Trials (CENTRAL) (the Cochrane Library, Issue 7). See Appendix 1.
- MEDLINE (via OvidSP) from 1996 to 3 August 2020. See Appendix 2.
- Embase (via OvidSP) from 1996 to 3 August 2020. See Appendix 3.
- The World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (apps.who.int/trialsearch) for all prospectively registered and ongoing trials. See Appendix 4.
- ClinicalTrials.gov (clinicaltrials.gov). See Appendix 5.

Searching other resources

- Bibliographic searching

We manually searched the reference lists of identified relevant trials or reviews for further studies.

We planned to obtain a copy of the full article for each record reporting a potentially eligible trial or study; where this was not possible or where relevant information was missing, we attempted to contact study authors for the additional information.

Data collection and analysis

Selection of studies

Two review authors (AB and EA) independently reviewed the titles and abstracts retrieved by the search to assess eligibility against inclusion criteria. We planned that if a review author had published a potentially eligible study, then two other review authors would review the study for eligibility; however, this did not occur. One review author (EA) obtained the full-text copies of all potentially relevant papers, and two review authors (AB and EA) reviewed the full texts. Any disagreements at this stage would be resolved by another review author (MT), and the included and excluded studies recorded. We planned to contact the authors of primary studies for clarification when needed. We recorded duplicate studies as one reference. We only included studies published in English. We recorded the selection process in a PRISMA flow diagram (Liberati 2009), in Review Manager 5 software (Review Manager 2020). We recorded the process using the Covidence system (Covidence).

Data extraction and management

Two review authors (AB and EA) independently completed data extraction using standardised custom forms designed to capture data for each included study separately. One review author (EA) created the forms, which another review author (AB) pilot tested. Both review authors then finalised the forms for use in the review, and one review author (EA) combined the forms from both review authors (AB and EA) into a unified summary. We planned that one review author (MT) would resolve any disagreements; however, this was not required. We planned to collect the following information.

- Study design: type of study, e.g. a nested case-control study from a randomised trial, or a non-randomised cohort study, or a case-control study. If there was matching, then what was matching by and to what level (e.g. age to plus/minus two years). Control group: yes/no (women without treatment). Whether prognostic or predictive study, or both. For prognostic factor study, what phase (following Altman 1998; Riley 2009).
- Participants: demographic information, including number of participants, age, body mass index (BMI), ethnicity, education. Summary statistics such as mean, interquartile range (or standard deviation) and range for age, BMI and absolute or relative baseline risk, or both, from a risk model (e.g. Gail model (Gail 1989), Tyrer-Cuzick (Tyrer 2004), Breast Cancer Surveillance Consortium (BCSC) (Tice 2008)). Total number and total number (percentage) postmenopausal, perimenopausal or premenopausal. For a predictive review, the previous variables were to be split by treatment or control group.
- Biomarker: whether mammograms were from film (digitised for density or not) or full-field digital mammography. Manufacturer of digital mammogram machine. Whether any preprocessing

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

11

![Cochrane Library logo] **Trusted evidence.
Informed decisions.
Better health.**

Cochrane Database of Systematic Reviews

was carried out for quality control of mammographic density. Density measure(s), and the range and average time between baseline mammogram and diagnosis, between diagnosis and start of endocrine therapy (or study entry), and between start of endocrine therapy (or study entry) and the follow-up mammogram.

- Setting: country, whether in a high-risk clinic, a treatment clinic, time period, urban/rural.
- Cointerventions: HRT use, chemotherapy use (treatment), targeted therapy use (treatment), radiotherapy use (treatment), neoadjuvant endocrine therapy use (treatment).
- Follow-up time period: minimum, mean, median, interquartile range, standard deviation, maximum follow-up.
- Sources of funding and stated conflicts of interest: descriptive text copied from sections in each paper.

We documented all results, including subgroup analyses. When publications pertained to more than one publication, we planned to extract the data from all publications and record them in the database as such. We planned to consider the most recent or up-to-date reference (largest number of participants, or longest follow-up time, or correction to previous analysis) as the primary reference.

Assessment of risk of bias in included studies

For the prognostic review, we used a version of the Quality in Prognostic Studies (QUIPS) tool (Hayden 2013), modified for our review (Table 1), to assess the risk of bias (Hayden 2006). We used this tool to assess six important domains that might affect bias in the included studies: (i) study participation, (ii) attrition, (iii) measurement of density, (iv) measurement of the outcomes, (v) confounding, and (vi) statistical analysis. Each domain was rated as 'high', 'moderate' or 'low' risk, as per the QUIPs tool (Hayden 2013). Assessment of 'high', 'moderate' or 'low' risk of bias was based on the prompting items outlined in Table 1.

For the predictive biomarker review, we augmented the QUIPS tool with the Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) tool (Table 2; Table 3) (Sterne 2016). We used this tool to assess the risk of bias in estimation of an interaction between mammographic density change and treatment. Two review authors (AB and EA) independently assessed risk of bias in the studies, and one review author (EA) combined the independent assessments into one unified assessment for each included study. We planned that another review author (MT) would resolve any disagreements. If a review author was an author of an included study, then two other review authors would independently complete data extraction and assess the study for risk of bias; however, this was not required.

For both prognostic and predictive biomarker reviews, we considered the included studies together but with a narrative identifying the risk in different domains across studies. We planned to exclude studies that had substantial potential for bias in a sensitivity analysis of results.

Measures of biomarker response

Effect measure

In both reviews, the primary measure sought was the mean effect over a five-year follow-up period. We planned to allow other time periods, but if split into different periods (e.g. 0 to 5 years; 5 to 10 years), periods outside the initial five years would be in a secondary analysis. We planned any quantitative meta-analysis results to be in subgroups by similar cutpoints and by those using continuous trends. We reported the ratios so that less than 1.0 favours a risk reduction associated with a decrease in mammographic density, and greater than 1.0 indicates a risk increase.

Prognostic biomarker review

The primary measure sought was a hazard ratio (cohort study with time to event) or an odds ratio (case-control study) for the effect of density change. We planned to treat an odds ratio as an equivalent measure of the hazard ratio, unless rates were high, in which case we would include the odds ratio estimates in a secondary analysis.

Predictive biomarker review

The primary measure sought was an interaction between treatment and the biomarker, expressed as a relative hazard (cohort study) or odds ratio (case-control study).

Adjustment

Prognostic biomarker review

The primary effect estimate sought was to be adjusted. We planned to include unadjusted estimates if adjusted estimates were not available. To measure the prognostic ability of factors, it is commonly accepted that effect estimates that are adjusted for potential confounders are more relevant than unadjusted ones (Riley 2013). However, we planned that when adjusted estimates were not available, we would use unadjusted estimates because we did not expect the change in density to be associated with the baseline value of most other prognostic factors. We acknowledge that changes in BMI may also have occurred, and since BMI is negatively associated with breast density and a prognostic factor, one would ideally adjust for this in the analysis.

Predictive biomarker review

The primary effect estimate was adjusted. There are currently no established predictive biomarkers for either prevention or treatment in the groups of women that we planned to include in the review as defined above.

Dealing with missing data

We planned that when important data were missing, we would contact study authors to obtain the data.

Assessment of heterogeneity

We planned to measure heterogeneity in meta-analysis results using the estimated variance in a random-effects model (Tau^2). We likewise planned to assess publication bias using a funnel plot and Egger's test (Egger 1997).

Subgroup analysis and investigation of heterogeneity

We planned that in the case of sufficient included studies we would conduct the following a priori subgroup analysis to explore reasons for heterogeneity within the predefined homogeneous groups above.

Between studies

- Drug within SERM (tamoxifen, raloxifene, lasofoxifene, arzoxifene, droloxifene, bazedoxifene, fulvestrant) and AI grouping (anastrozole, letrozole, exemestane).
- Type of study: case-control, observational cohort, randomised trial (nested case-control).
- Type of cancer at baseline (treatment): (percentage DCIS).
- Severity of cancer at baseline (treatment): stage (percentage regional spread).
- Co-interventions (treatment): chemotherapy/targeted therapy.
- Hormone therapy use during therapy (yes/no, percentage if available), or in previous two years (yes/no, percentage if available).
- Time between start of therapy (or study entry) and follow-up mammogram (mean and range).
- Menopausal status (percentage premenopausal).
- Age (mean).
- BMI (mean).
- Digital or film mammography (percentage digital).
- Distribution of density at baseline (some studies may exclude women with low density).

Within-study estimates of effect

- Type of cancer at baseline (treatment): DCIS versus invasive.
- Severity of cancer at baseline (treatment): stage (percentage regional spread).
- Co-interventions (treatment): chemotherapy/targeted therapy.
- Hormone therapy use: no HRT prior to endocrine therapy; some HRT two years or more than two years prior to endocrine therapy; some HRT less than two years prior to endocrine therapy; some HRT during endocrine therapy.
- Menopausal status (pre-, peri-, or postmenopausal).
- Age group (< 50 years or ≥ 50 years) as a proxy for menopausal status.
- BMI (both within-study (< 25, 25 to < 30, 30 to < 35, ≥ 35 kg/m^2) and between studies (mean)).
- Baseline density.

Data synthesis

We expected heterogeneity between studies in this review as it is common in reviews of prognostic biomarkers (Riley 2013). To address this we planned to only consider undertaking meta-analysis for studies within predefined groups considered homogenous enough in advance to be meaningful for data synthesis, namely those with the same class of drug, same outcome, same density measure, same effect measure (same cutpoint or continuous variable assessment). Where more than one study was available, we planned to combine estimates using an inverse-variance weighting (fixed-effect estimation); if there was substantial variability, then we would present the result but state that the overall effect estimate has very limited interpretation, whilst we planned to seek subgroups (see above) that best explained the heterogeneity.

Meta-analysis of the studies using individual data from participants may have overcome many of the expected issues arising in this review of published data, including heterogeneity in the biomarker used and cutpoints (Riley 2009; Riley 2013). In future work, we plan to use the review to identify relevant studies, and pool data from the best-quality studies (using information from the risk of bias analysis) in an individual participant-level review.

R E S U L T S

Description of studies

Results of the search

The database search identified 1264 records (see PRISMA flow diagram in Figure 1); after removal of duplicates, 970 records remained. Of these, 878 records were deemed ineligible based on title and abstract, and 92 records were selected for full-text review. We excluded 84 full-text articles, and included eight eligible studies that fulfilled the inclusion criteria in the qualitative synthesis, of which seven provided data on endpoints listed in our protocol. We also conducted a bibliography search of the reference lists of the eight included studies. We reviewed titles; for those records we considered potentially eligible (and not already included in the 970 records), we also reviewed abstracts of the studies. We identified nine potentially relevant records, but considered these to be ineligible after abstract review, hence we found no additional studies through the bibliographic search.

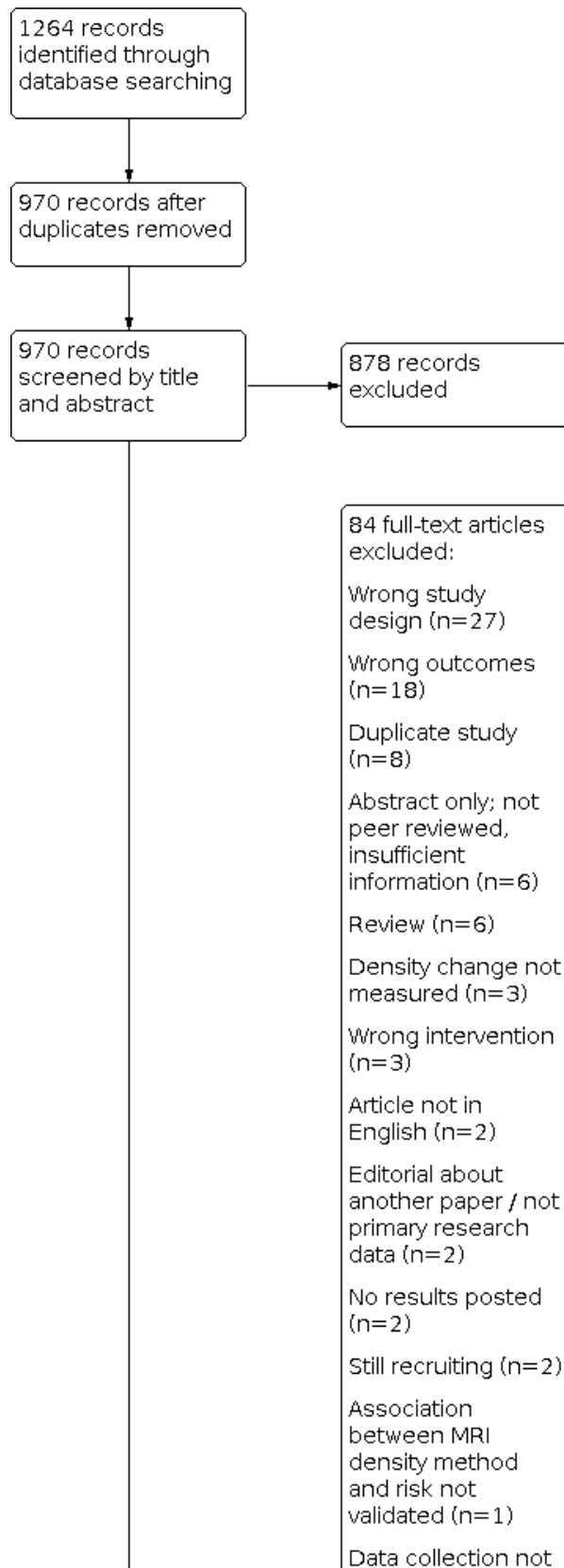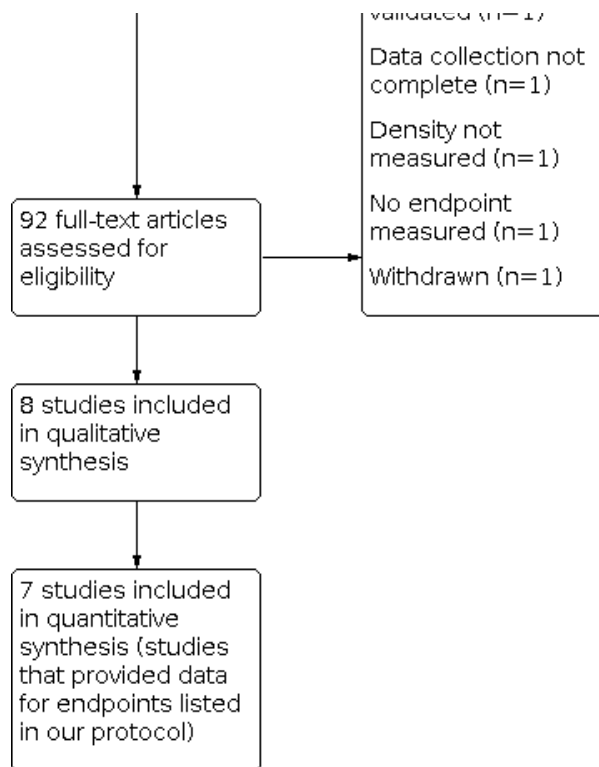**Figure 1. Study flow diagram.**

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

14

Figure 1. (Continued)

```
                    ┌─────────────────┐        ┌──────────────────────┐
                    │ 92 full-text    │        │ validated (n=1)      │
                    │ articles        │        │                      │
                    │ assessed for    │───────▶│ Data collection not  │
                    │ eligibility     │        │ complete (n=1)       │
                    └─────────────────┘        │                      │
                            │                  │ Density not          │
                            │                  │ measured (n=1)       │
                            │                  │                      │
                            │                  │ No endpoint          │
                            │                  │ measured (n=1)       │
                            │                  │                      │
                            │                  │ Withdrawn (n=1)      │
                            ▼                  └──────────────────────┘
                    ┌─────────────────┐
                    │ 8 studies       │
                    │ included        │
                    │ in qualitative  │
                    │ synthesis       │
                    └─────────────────┘
                            │
                            ▼
                    ┌─────────────────┐
                    │ 7 studies       │
                    │ included        │
                    │ in quantitative │
                    │ synthesis       │
                    │ (studies        │
                    │ that provided   │
                    │ data            │
                    │ for endpoints   │
                    │ listed in our   │
                    │ protocol)       │
                    └─────────────────┘
```

Included studies

Study design, sample size, country of recruitment, and follow-up period

There was a large amount of variation across the eight included studies (see Characteristics of included studies section). Two studies included data from randomised controlled trials (Cuzick 2011a; van Nes 2015), and there were four case-control studies, Cuzick 2011a; Knight 2018; Nyante 2015; Sandberg 2013, and four cohort studies with participants identified retrospectively (Kim 2012; Ko 2013; Li 2013; van Nes 2015). Three of the four case-control studies used matching (Knight 2018; Nyante 2015; Sandberg 2013). The studies ranged in size from 349 women, in Nyante 2015, to 1066 women, in Ko 2013. Six studies included women from Western populations (Canada (Knight 2018), the UK and Finland (Cuzick 2011a), the Netherlands (van Nes 2015), Sweden (Li 2013; Sandberg 2013), and the USA (Knight 2018; Nyante 2015)); two studies were conducted in women from South Korea (Kim 2012; Ko 2013). Maximum follow-up ranged from 5 years, in Ko 2013, to 14 years, in Li 2013.

Setting and type of endocrine therapy

Only one study was conducted in the preventive setting (Cuzick 2011a), with the rest being performed in the adjuvant treatment setting. Half of the included studies assessed tamoxifen treatment only (Cuzick 2011a; Ko 2013; Li 2013; Nyante 2015); two studies assessed tamoxifen and an AI (Kim 2012; van Nes 2015); and two studies included a subset of women who were identified as being on endocrine therapy during their adjuvant treatment (Knight 2018; Sandberg 2013). Two studies included a placebo, Cuzick 2011a, or control group (Li 2013), but only Cuzick 2011a was eligible for the predictive biomarker review. The included studies involved pre-

and postmenopausal women. In the treatment setting, two South Korean studies included women with DCIS or invasive breast cancer at entry (Kim 2012; Ko 2013), whereas the other treatment studies included women with invasive disease only (Knight 2018; Li 2013; Nyante 2015; Sandberg 2013; van Nes 2015).

Breast density definitions

The two studies from clinical trials used visually assessed density, with one assessing density to the nearest 5% (Cuzick 2011a), and the other assessing density in Boyd categories (van Nes 2015). Two studies used a machine-learning-based density assessment that was designed to automate breast density assessment by an expert using a semi-automated thresholding algorithm (Cumulus) (Li 2013; Sandberg 2013). Three studies used Cumulus per cent density (Kim 2012; Knight 2018; Nyante 2015), and one study used BI-RADS density (Ko 2013). Density change cutpoints varied. Some adopted an absolute change in percentage density of 5%, Kim 2012, or 10% (Cuzick 2011a; Knight 2018; Sandberg 2013). One study used an absolute per cent density reduction cutpoint based on tertiles determined by the distribution of controls (Nyante 2015). Another study used a 20% relative dense area reduction cutpoint (Li 2013). A further approach was to record any reduction in BI-RADS category (Ko 2013). One study did not include a cutpoint, but instead modelled the effect of density change as a per-unit reduction, where each Boyd category was treated as a unit (van Nes 2015).

Outcomes assessed

Two studies reported results for the recurrence endpoint (recurrence-free survival) (Kim 2012; Ko 2013); two studies reported mortality (Li 2013; Nyante 2015); two studies reported incidence of contralateral breast cancer (Knight 2018; Sandberg 2013);

one study assessed incidence of contralateral breast cancer and recurrence (loco-regional and distance) (van Nes 2015); and the final study assessed risk of developing a first invasive or DCIS breast cancer (Cuzick 2011a). A detailed description of the included studies is provided in the Characteristics of included studies section.

Risk of bias in included studies

The risk of bias in each domain and by each study is outlined in Figure 2. In Figure 2, high risk is presented as '-', moderate risk as '?' and low risk of bias as '+'. A detailed description of the risk of bias judgements is provided in the Characteristics of included studies section. A summary by domain is provided below.

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias domain for each included study; "+": low risk of bias, "?": moderate risk of bias, "-": high risk of bias.

	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting
Cuzick 2011a	+	+	?	+	+	+
Kim 2012	?	?	?	?	-	-
Knight 2018	-	+	?	+	?	-
Ko 2013	?	?	?	?	?	-
Li 2013	?	+	?	+	+	+
Nyante 2015	?	+	+	+	+	+
Sandberg 2013	?	+	+	+	?	?
van Nes 2015	+	+	+	+	-	?

Study participation

The risk of bias for study participation was moderate for most studies due to the risk that the study samples might not reflect the target population of this review. One issue was that many women from the wider study sample were often excluded due to mammogram availability. Only two studies from clinical trials conducted analysis of differences in prognostic and confounding factors between those women included and excluded due to breast density change availability (Cuzick 2011a; van Nes 2015); these had the lowest risk of bias for study participation. We assessed one study as at high risk of bias in part because it required women to still be alive at the time of interview for the study, therefore it lacked

information on women who had died before recruitment, and so might exclude information on those with the poorest prognosis (Knight 2018).

Study attrition

We assessed most studies as at low risk, but rated two studies as moderate due to a lack of information on participant dropout, loss to follow-up, reasons for censoring, or adherence to endocrine therapy (Kim 2012; Ko 2013).

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review) 16

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Prognostic factor measurement

We scored the risk of bias for the measurement of mammographic density as moderate or low. We deemed studies to be at a moderate risk of bias if they did not provide information on aspects considered in the review protocol regarding accuracy of the measure, such as the experience of the reader(s), the time between baseline and follow-up mammogram, blinding to treatment (when appropriate) or case-control status, reliability of measurements, number of women who had density measured, or reasons for missing density data. In addition, we assessed any study that did not use a predefined cutpoint for density change as at moderate risk of bias for this domain.

Outcome measurement

The domain outcome measurement was at the lowest risk of bias overall. Outcomes were based on clinical trial databases (Cuzick 2011a; van Nes 2015), population-based registries (Knight 2018; Li 2013; Nyante 2015; Sandberg 2013), or hospital records (Kim 2012; Ko 2013). It was not clear how accurate or complete information from the hospital records was, therefore we assessed the last two studies as moderate (Kim 2012; Ko 2013).

Study confounding

Risk of bias due to confounding was mixed across studies. Adjustments were unclear in two studies (Kim 2012; Sandberg 2013), and no adjustments were made in another study (van Nes 2015). Several studies did not adjust for chemotherapy (despite the knowledge that it can induce menopause in premenopausal women, causing an oestrogen deprivation and reduction in density, as well as reduced risk of recurrence) (Kim 2012; Ko 2013; Nyante 2015; Sandberg 2013; van Nes 2015). Only four studies considered adherence to therapy (tamoxifen reduces breast density and is an effective treatment, therefore density change could be confounded by adherence) (Cuzick 2011a; Li 2013; Nyante 2015; van Nes 2015).

Statistical analysis and reporting

The domain statistical analysis and reporting was at the highest risk of bias overall. The statistical methods for survival analysis were not adequately defined in two studies (Kim 2012; Ko 2013). Additionally, two studies did not fully describe the data for the subgroups of interest, and results were not reported by individual endocrine therapy (Sandberg 2013; van Nes 2015).

Findings

There were very few eligible studies for each outcome. We deemed heterogeneity across mammographic density measures and effect measures (including cutpoints for density change) as too great to conduct a meaningful meta-analysis. Instead, the findings for each outcome are described below and are further summarised in the Summary of findings 1 (prognostic biomarker - treatment studies), Summary of findings 2 (prognostic biomarker - prevention studies), and Summary of findings 3 (predictive biomarker).

Firstly, the results for the prognostic biomarker are described.

Prognostic biomarker

Primary outcomes

Potential benefits from treatment

Treatment: breast cancer mortality (time to death caused by breast cancer)

Two studies assessed density change as a prognostic biomarker for breast cancer mortality with tamoxifen treatment (Li 2013; Nyante 2015). A total of 823 women were included (172 deaths from breast cancer: 97 cases from a case-control study; 75 breast cancer deaths during follow-up from a cohort study). Mammographic density was assessed two ways (Cumulus per cent density from mammograms 3 to 26 months apart, and an automated machine-learning measure from mammograms 6 to 36 months apart). A reduction in density reflected 56% and 50% reductions in breast cancer mortality (odds ratio (OR) 0.44 (95% confidence interval (CI) 0.22 to 0.88) for Cumulus per cent density reduction > 8.7% compared with < 0.5%; hazard ratio (HR) 0.50 (95% CI 0.27 to 0.93) for relative reduction in automated dense area > 20% compared with ≤ 9% increase to ≤ 10% reduction). We assessed the certainty of the evidence for breast cancer mortality as low.

Prevention: incidence of invasive breast cancer and DCIS

One study assessed density change as a prognostic biomarker for the incidence of invasive breast cancer and DCIS (Cuzick 2011a). The nested case-control study (from within a multicentre international randomised controlled trial) included 51 women who subsequently developed breast cancer and 456 women who were breast cancer free at the time the study was conducted, all of whom were on preventive tamoxifen. Per cent density was assessed visually in 5% bands from mammograms 12 to 18 months apart. A reduction in density reflected a 68% reduction in incidence of invasive breast cancer and DCIS (OR 0.32 (95% CI 0.14 to 0.72) for ≥ 10% reduction in visual per cent density compared with 0% density change). We assessed the certainty of the evidence for incidence of invasive breast cancer and DCIS (prognostic biomarker) as low.

Potential harms from treatment

Treatment and prevention: rate of all serious adverse events. These included serious side effects noted for tamoxifen (cataracts, pulmonary embolism or deep vein thrombosis, and endometrial cancer) and anastrozole (osteoporosis and bone fractures).

These outcomes were not collected or reported in the included studies.

Secondary outcomes

Potential benefits from treatment

Treatment: recurrence

Three studies assessed density change as a prognostic biomarker for breast cancer recurrence with endocrine treatment (Kim 2012; Ko 2013; van Nes 2015).

A total of 1956 women were included for tamoxifen treatment (the number of breast cancer recurrences during follow-up in two cohort studies was unclear, but we calculated a maximum of 147 events). Mammographic density was assessed two ways (Cumulus per cent density from mammograms 8 to 20 months apart, and BI-RADS density on mammograms approximately 10 to 34 months apart). A reduction in density reflected 34% and 65% reductions in breast cancer recurrence, although these point

estimates (most likely value) were not precise: the confidence interval for one of the studies included the null effect (HR 0.66 (95% CI 0.40 to 1.09) for Cumulus per cent density reduction ≥ 5% compared with < 5%; HR 0.35 (95% CI 0.17 to 0.68) for reduction in BI-RADS density categories compared with no reduction). There was considerable risk of bias due to study participation, attrition, prognostic factor measurement, outcome measurement, and statistical methodology (there were unclear adjustments and a potential immortal time bias).

A total of 175 women were included for aromatase inhibitor treatment (the number of breast cancer recurrences during follow-up in one cohort study was unclear, but we calculated a maximum of 80 events). Mammographic density was assessed using Cumulus per cent density from mammograms 8 to 20 months apart, and a reduction in density reflected an 86% reduction in breast cancer recurrence, although this effect was not precise as the confidence interval included the null effect (HR 0.14 (95% CI 0.02 to 1.11) for Cumulus per cent density reduction ≥ 5% compared with < 5%). There was considerable risk of bias due to study participation, attrition, prognostic factor measurement, outcome measurement, and statistical methodology (there were unclear adjustments and a potential immortal time bias).

One cohort study (within a randomised controlled trial) also reported results for women on endocrine treatment (exemestane or tamoxifen); however, these were not separated by individual drug. Recurrence was also reported separately for location of recurrence (loco-regional or distance). A total of 284 women were included in both the loco-regional and distance recurrence analysis (the number of breast cancer recurrences diagnosed during follow-up was unreported), and mammographic density was assessed visually using Boyd categories from mammograms 18 to 30 months apart. Results from this study of unspecified endocrine therapies were inconsistent with the other included studies assessing the recurrence outcome. A reduction in density reflected a 48% and 32% increase in breast cancer recurrence for loco-regional and distance recurrence, respectively. However, these effects were not precise, as the confidence intervals were wide and included the null effect (HR for loco-regional 1.48 (95% CI 0.38 to 5.81) per Boyd category reduction; HR for distance 1.32 (95% CI 0.64 to 2.71) per Boyd category reduction). The analyses did not adjust for important confounding factors such as age, BMI, baseline density, time between mammograms, and chemotherapy.

We assessed the certainty of the evidence for all recurrence endpoints as very low.

Treatment: incidence of a secondary primary breast cancer (e.g. in the contralateral breast)

Two studies (one from a randomised controlled trial, one retrospective observational study) assessed density change as a prognostic biomarker for the incidence of a secondary primary breast cancer (Sandberg 2013; van Nes 2015). Results were reported for women treated with endocrine therapy (exemestane, tamoxifen, or unknown); however, these were not separated by individual drug. A total of 451 women were included (the number of secondary primary breast cancers was unreported in both studies, but we calculated a maximum of 364 events). Mammographic density was assessed two ways (an automated machine-learning measure on mammograms 12 to 60 months apart, and visual assessment of Boyd categories from mammograms 18 to 30 months

apart). A reduction in density reflected 48% and 42% reductions in the incidence of a contralateral breast cancer, although this effect was not precise, as the confidence intervals were wide and included the null effect (OR 0.52 (95% CI 0.18 to 1.51) for automated per cent density reduction ≥ 10% compared with < 10% increase to < 10% reduction; HR 0.58 (95% CI 0.08 to 4.44) per Boyd category reduction). The evidence for this outcome was indirect as the analysis of interest in both studies was a secondary objective, therefore there was a lack of reporting on data for subgroups, individual endocrine therapies and adjustments used in the analysis. We assessed the certainty of the evidence for incidence of a secondary primary breast cancer as very low.

Treatment: any recurrence or any death (disease-free survival)

This outcome was not collected or reported in the included studies.

Treatment: distant metastases

This outcome was not collected or reported in the included studies.

Treatment: death from all causes (all-cause mortality)

This outcome was not collected or reported in the included studies.

Treatment: recurrence of invasive cancer only

This outcome was not collected or reported in the included studies.

Treatment: recurrence of DCIS cancer only

This outcome was not collected or reported in the included studies.

Prevention: incidence of invasive cancer only

This outcome was not collected or reported in the included studies.

Prevention: incidence of DCIS cancer only

This outcome was not collected or reported in the included studies.

Potential harms from treatment

Treatment and prevention: troublesome but not serious side effects observed for SERMs and AIs, including vasomotor symptoms and joint or muscle pain.

These outcomes were not collected or reported in the included studies.

Next, we describe the results for the predictive biomarker.

Predictive biomarker

Primary outcomes

Potential benefits from treatment

Treatment: breast cancer mortality (time to death caused by breast cancer)

This outcome was not collected or reported in the included studies.

Prevention: incidence of invasive breast cancer and DCIS

Only one study evaluated density change as a predictive biomarker (Cuzick 2011a). The nested case-control study (from within a multicentre, international randomised controlled trial) included 123 women with breast cancer and 942 women without breast cancer (51 subsequent breast cancers for women who were on preventive tamoxifen). Per cent density was assessed visually in 5% bands from mammograms 12 to 18 months apart. The study

reported an OR of 0.53 (95% CI 0.21 to 1.32) for an interaction between visual per cent density reduction (≥ 10% or < 10%) and prophylactic tamoxifen, although this point estimate was not precise, as the confidence interval was wide and included the null effect. The power to detect an interaction was limited by the sample size and the number of cases in the study. We assessed the certainty of the evidence for incidence of invasive breast cancer and DCIS (predictive biomarker) as low.

Potential harms from treatment

Treatment and prevention: rate of all serious adverse events. These included serious side effects noted for tamoxifen (cataracts, pulmonary embolism or deep vein thrombosis, and endometrial cancer) and anastrozole (osteoporosis and bone fractures).

These outcomes were not collected or reported in the included studies.

Secondary outcomes

The remaining outcomes for the predictive biomarker (equivalent to those listed above for the prognostic biomarker) were not collected or reported in the included studies.

Eight studies (including Knight 2018) contributed to the prespecified outcomes; however, seven studies contributed data (Knight 2018 did not report data in a format amenable for data extraction).

DISCUSSION

Summary of main results

This review considered evidence that mammographic density is a prognostic or predictive biomarker in women receiving endocrine therapy. Most studies on prognosis for women receiving endocrine therapy reported a reduced risk associated with breast density change. Across endpoints, settings, and drugs, risk ratio point estimates (most likely value) were between 0.1 and 1.5, but with substantial uncertainty. In particular, the data suggest that breast density change is a prognostic biomarker in women receiving tamoxifen for treatment or prevention, potentially with a large effect (across endpoints and settings, prognostic risk ratio point estimates for tamoxifen were between 0.3 and 0.7). There was some evidence to suggest that breast density change is a predictive biomarker in women receiving tamoxifen (a nested case-control study from a randomised breast cancer tamoxifen prevention trial reported a point estimate of 0.5, but with wide uncertainty on the estimated interaction). Overall, the level of evidence available was limited by the number of studies, as well as risk of bias in the included studies.

Most of the included studies focused on the mammographic density biomarker in women receiving tamoxifen, whilst only Kim 2012 provided results for the density change biomarker for treatment with AIs. Sandberg 2013 included women on endocrine therapy but did not stipulate which drug(s), and whilst van Nes 2015 included women on both tamoxifen and exemestane, the results were not reported separately by each drug. Only Cuzick 2011a contributed to the predictive biomarker review, whereas seven studies were included in the prognostic biomarker review (Cuzick 2011a; Kim 2012; Ko 2013; Li 2013; Nyante 2015; Sandberg 2013; van Nes 2015). Similarly, only Cuzick 2011a was conducted in the preventive setting, whilst six studies were in the adjuvant

setting (Kim 2012; Ko 2013; Li 2013; Nyante 2015; Sandberg 2013; van Nes 2015). We note that the outcome for incidence of a contralateral breast cancer reported in Sandberg 2013 and van Nes 2015 has implications for prevention, since a reduction in mammographic density after endocrine therapy could be indicative of a reduction in risk of a secondary primary breast cancer. However, Sandberg 2013 and van Nes 2015 were not considered for inclusion in the preventive setting because all women had been diagnosed with a first breast cancer at entry to the study.

There was a great deal of variation in definition of breast density change between studies. Density measurements included visual percentage score, visual percentage bands (Boyd categories), BI-RADS categorical assessment, Cumulus semi-automated percentage score and fully automated percentage and absolute scores. Most studies used a cutpoint for density reduction, including 5%, 10% or tertile cutpoints (absolute percentage density reduction), a 20% cutpoint (relative dense area reduction), and BI-RADS categories. The characteristics of participants also varied. Studies were conducted in European, North American, and Asian populations of premenopausal and postmenopausal women, with different disease status at first diagnosis in treatment studies.

Our assessment of the certainty of the evidence was very low or low, as shown in Summary of findings 1; Summary of findings 2; Summary of findings 3. This was mainly due to the small number of included studies, with all having an observational design and high or moderate risk of bias in at least one risk domain.

Applicability of findings to clinical practice and policy

The evidence in this review is limited, and more studies are required to improve the certainty of evidence that mammographic density is a prognostic or predictive biomarker for women receiving endocrine therapy. Our findings suggest that the level of evidence is low, very low, or not available, and further research is warranted for all review endpoints and classes of endocrine therapy.

Quality of the evidence

There was a large amount of heterogeneity across studies, with only seven studies included in the synthesis (Cuzick 2011a; Kim 2012; Ko 2013; Li 2013; Nyante 2015; Sandberg 2013; van Nes 2015). The evidence for most outcomes was initially graded as low certainty due to the observational design of the contributing studies (Schünemann 2013). The evidence for the outcomes 'recurrence (loco-regional)', 'recurrence (distance)', and 'incidence of invasive breast cancer and DCIS' was initially graded as moderate certainty as the data for these outcomes were from randomised controlled trials. The level of the evidence was upgraded or downgraded in accordance with GRADE guidelines (Schünemann 2011; Schünemann 2013). We assessed the evidence for the recurrence (overall, loco-regional, and distance) and contralateral breast cancer outcomes as of very low certainty due to high risk of bias and imprecision. Additionally, the recurrence outcomes (with AI treatment, loco-regional and distance) had evidence from only one study each (Kim 2012; van Nes 2015). We assessed the evidence for the outcomes 'incidence of invasive breast cancer and DCIS' and 'breast cancer mortality' as of low certainty. The influence of bias was relatively low for these outcomes, but this was counterbalanced by the small number of contributing studies; for instance, breast cancer incidence was assessed in only one study for both the prognostic and predictive biomarker reviews

(Cuzick 2011a). We downgraded the certainty of the evidence from moderate to low for the outcome incidence of breast cancer in the prognostic biomarker review because the mammographic density cutpoint was determined by the observed data in the one contributing study. We downgraded the certainty of the evidence for the outcome incidence of breast cancer for the predictive biomarker review because the one contributing study reported an imprecise effect with a confidence interval covering the null.

Strengths and weaknesses of the review

A potential bias of the review process is the exclusion of grey literature, which included six conference abstracts (Cosmacini 1993; Decensi 2004; Macis 2011; Martin 2016; Ozhand 2013; Redfern 2016b). We regarded these as ineligible because they had not been peer reviewed, and the abstracts were limited in the amount of information reported, such as full statistical methodology, justification for chosen techniques, and number of women with density reductions by outcome. These studies may be eligible for inclusion in a subsequent updated version of this review once they are reported as peer-reviewed full texts. Similarly, one study did not report any results (NCT00316836), and two studies were still recruiting (NCT00445445; NCT01765049); these studies may also be eligible for inclusion in an updated review once results are published as peer-reviewed full texts. One study used a non-validated magnetic resonance imaging (MRI) density method (Kim 2014), which may be eligible in an updated review if more than one study (outside of the review studies) can show an association between the density method and breast cancer risk.

Furthermore, to determine how many other studies would have been included in the review had language restrictions not been in place, we conducted a post hoc search of Google Scholar and PubMed to identify non-English papers citing any of the eight studies included in this review (Google Scholar; PubMed). We found 29 non-English citations to at least one of the eight included studies in Google Scholar. To assess potential eligibility, the records were translated into English using Google Translate (Google Translate). Assessment of the abstracts (by EA) concluded that all 29 records would have been ineligible for inclusion (records had inappropriate study design (n = 8), were reviews of other articles, hence did not include primary research data (n = 10), or were theses which had not been published as peer-reviewed full-text articles (n = 11)). We did not find any non-English records through our post hoc search of PubMed (PubMed), but from the main search excluded two records at the full-text screening stage because they were not in English (Boisserie-Lacroix 2008; Boutet 2004). These two records were also ineligible, as they were reviews of other articles and did not include primary research data. We therefore determined that the risk of bias from missing non-English records in our search for studies was low.

Finally, we identified two studies with the same study participants (Mullooly 2019; Nyante 2015). We chose to include Nyante 2015 over Mullooly 2019 because the effect of mammographic density change on breast cancer-specific death was only a secondary objective in Mullooly 2019, therefore their analysis included a subset of the participants from Nyante 2015, whereas Nyante 2015 evaluated the effect of mammographic density change on breast cancer mortality as a primary objective and hence included the full set of participants. We therefore judged the risk of bias from excluding Mullooly 2019 to be low.

Agreement and disagreements with other studies or reviews

An earlier systematic review by Shawky 2017 reported findings similar to our review, although an assessment of quality was not included in their paper. Their analysis included a study excluded from our review because it was a conference abstract (Redfern 2016b). The study by Knight 2018, which was included in our review, was published after Shawky 2017. Another recent systematic review, Kanbayti 2019, also reported similar findings to our review, although their review assessed the relationship between mammographic density reduction and patient outcomes in women receiving any type of breast cancer treatment, therefore the focus was not specific to endocrine therapy. Furthermore, Kanbayti 2019 included two studies that were excluded at the full-text screening stage of our review (Andersson 2017; Kim 2014). This was due to differences in inclusion criteria: the study by Kanbayti 2019 did not specify that the density measurement was to have been validated as a breast cancer risk factor in more than one study (outside of the review studies), and the density biomarker did not have to be derived from a baseline and a follow-up mammogram (as was stipulated in our review). Moreover, Kanbayti 2019 did not include the study by Cuzick 2011a because their population of interest was breast cancer patients treated in the adjuvant setting only.

A U T H O R S ' C O N C L U S I O N S

Implications for practice

There is low-/very low-level evidence to support using mammographic density reductions following endocrine therapy as a prognostic biomarker for treatment or prevention. There is low-level evidence to support them as a predictive biomarker in prevention, and no evidence at all in the treatment setting.

Implications for research

More research is warranted to assess whether mammographic density reductions associated with endocrine therapy are prognostic or predictive biomarkers. The level of evidence is low or very low for all endpoints. We identified no studies for several of the secondary outcomes, and also for potential harms.

When conducting this review we identified a number of shortcomings in the current literature that will be important to consider for those conducting more primary research in this area. We end by discussing some of these.

Adherence to endocrine therapy is a confounding effect on the relationship between endocrine therapy-induced density change and breast cancer outcome. Women who receive less treatment are expected to have smaller declines in breast density. They will also get less clinical benefit from the drug. Greater effort should be made to measure and control for the potential confounding effect of adherence in future studies.

Efforts should also be made to ensure (and report) that there has been no previous treatment with selective oestrogen receptor modulators (SERMs) or aromatase inhibitors (AIs) before women enter the study (e.g. for ductal carcinoma in situ or prevention). Residual effects of these previous treatments might influence density change and outcome during the study.

We included only one study for density change as a predictive biomarker. In future studies it will be important to define and identify a suitable control group. This group should be as similar to the treatment group as possible (ideally through randomisation), but will not receive endocrine therapy. This will be a difficult task in many settings, and we note that this is practically impossible in the adjuvant setting. However, we also note that control groups defined simply as "no tamoxifen" (as in Li 2013) are problematic, as they may have been treated with another SERM or AI, or belong to a different target population (e.g. oestrogen receptor-negative breast cancers); this should be avoided.

Furthermore, studies need to ensure that they use (and report) appropriate definitions of start and end of follow-up time. Follow-up time should start after the second mammogram, because women are not at risk before this time. Likewise, if a study stipulates that women must be on treatment for a certain amount of time (e.g. one year), then women are only at risk from the latest date of one year from start of treatment or follow-up mammogram, and follow-up should start from that point in time. Including the immortal time between mammograms in the analysis may bias findings towards showing a greater effect of density change on outcomes, given that there is evidence that a longer time between mammograms in women receiving tamoxifen is associated with greater declines in breast density; and would add greater immortal time into the risk estimate for women with larger density declines.

Finally, we found high heterogeneity between studies, particularly for the mammographic density biomarker and chosen cutpoints. One way of overcoming this issue is to pool data in a meta-analysis of individual participant-level data. However, we note that this might be challenging for our review question, due to such issues as those discussed above regarding timing of mammograms in relation to treatment and the large variety of different methods used to measure mammographic density.

ACKNOWLEDGEMENTS

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

21

REFERENCES

References to studies included in this review

Cuzick 2011a *{published and unpublished data}*

Cuzick J, Warwick J, Pinney E, Duffy S, Simon C, Howell A, et al. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case–control study. *Journal of the National Cancer Institute* 2011;**103**(9):744-52.

Kim 2012 *{published data only}*

Kim J, Han W, Moon HG, Ahn SK, Shin HC, You JM, et al. Breast density change as a predictive surrogate for response to adjuvant endocrine therapy in hormone receptor positive breast cancer. *Breast Cancer Research* 2012;**14**(4):R102.

Knight 2018 *{published data only (unpublished sought but not used)}*

Knight JA, Blackmore KM, Fan J, Malone KE, John EM, Lynch CF, et al. The association of mammographic density with risk of contralateral breast cancer and change in density with treatment in the WECARE study. *Breast Cancer Research* 2018;**20**(1):23.

Ko 2013 *{published data only}*

Ko KL, Shin IS, You JY, Jung SY, Ro J, Lee ES. Adjuvant tamoxifen-induced mammographic breast density reduction as a predictor for recurrence in estrogen receptor-positive premenopausal breast cancer patients. *Breast Cancer Research and Treatment* 2013;**142**(3):559-67.

Li 2013 *{published data only}*

Li J, Humphreys K, Eriksson L, Edgren G, Czene K, Hall P. Mammographic density reduction is a prognostic marker of response to adjuvant tamoxifen therapy in postmenopausal patients with breast cancer. *Journal of Clinical Oncology* 2013;**31**(18):2249-56.

Nyante 2015 *{published data only}*

Nyante SJ, Sherman ME, Pfeiffer RM, Berrington de Gonzalez A, Brinton LA, Aiello Bowles EJ, et al. Prognostic significance of mammographic density change after initiation of tamoxifen for ER-positive breast cancer. *Journal of the National Cancer Institute* 2015;**107**(3):dju425.

Sandberg 2013 *{published data only}*

Sandberg ME, Li J, Hall P, Hartman M, dos Santos Silva I, Humphreys K, et al. Change of mammographic density predicts the risk of contralateral breast cancer - a case-control study. *Breast Cancer Research* 2013;**15**(4):R57.

van Nes 2015 *{published and unpublished data}*

van Nes JG, Beex LV, Seynaeve C, Putter H, Sramek A, Lardenoije S, et al. Minimal impact of adjuvant exemestane or tamoxifen treatment on mammographic breast density in postmenopausal breast cancer patients: a Dutch TEAM trial analysis. *Acta Oncologica* 2015;**54**(3):349-60.

References to studies excluded from this review

Andersson 2017 *{published data only}*

Andersson TM, Crowther MJ, Czene K, Hall P, Humphreys K. Mammographic density reduction as a prognostic marker for postmenopausal breast cancer: results using a joint longitudinal-survival modelling approach. *American Journal of Epidemiology* 2017;**186**(9):1065-73.

Atkinson 1999 *{published data only}*

Atkinson C, Warren R, Bingham SA, Day NE. Mammographic patterns as a predictive biomarker of breast cancer risk: effect of tamoxifen. *Cancer Epidemiology, Biomarkers & Prevention* 1999;**8**:863–6.

Becker 2009 *{published data only}*

Becker S, Kaaks R. Exogenous and endogenous hormones, mammographic density and breast cancer risk: can mammographic density be considered an intermediate marker of risk? *Recent Results in Cancer Research* 2009;**181**:135-57.

Boisserie-Lacroix 2008 *{published data only}*

Boisserie-Lacroix M, Lebiez-Michel N, Cavagni P, Bentolila J, Laumonier H, Bouzgarrou M, et al. Hormones and mammographic breast density [Hormones et densité mammaire]. *Journal de Radiologie* 2008;**89**(9):1196-203.

Boutet 2004 *{published data only}*

Boutet G, Boisserie-Lacroix M, Trillaud H. Menopausal hormonal therapies: impact on mammographic breast density [Thérapeutiques hormonales de la ménopause: impact sur la densité mammographique]. *Journal de Radiologie* 2004;**85**(10 Pt 1):1673-86.

Boyd 2001 *{published data only}*

Boyd NF, Martin LJ, Stone J, Greenberg C, Minkin S, Yaffe MJ. Mammographic densities as a marker of human breast cancer risk and their use in chemoprevention. *Current Oncology Reports* 2001;**3**:314–21.

Boyd 2011 *{published data only}*

Boyd NF. Tamoxifen, mammographic density, and breast cancer prevention. *Journal of the National Cancer Institute* 2011;**103**(9):704-5.

Chlebowski 2003 *{published data only}*

Chlebowski RT, McTiernan A. Biological significance of interventions that change breast density. *Journal of the National Cancer Institute* 2003;**95**(1):4-5.

Cosmacini 1993 *{published data only}*

Cosmacini P, Cassano E, Sacchini V, Coopmans de Yoldi G, Costa A. Chemoprevention of breast cancer by using tamoxifen: radiologic evaluation. *Radiology* 1993;**189**(Suppl):245.

Cuzick 2012 *{published data only}*

Cuzick J. Breast density predicts endocrine treatment outcome in the adjuvant setting. *Breast Cancer Research* 2012;**14**(4):109.

Decensi 2004 *{published data only}*

Decensi A, Bonanni B, Guerrieri GA, Robertson C, Cazzaniga M, Mariette F, et al. A randomized 2x2 biomarker trial of low-dose tamoxifen and fenretinide in postmenopausal women at high-risk for breast cancer. In: American Society of Clinical Oncology Annual Meeting; 2004 June 5–8; New Orleans. Baltimore: Lippincott Williams & Wilkins, 2004:97.

Decensi 2009 *{published data only}*

Decensi A, Robertson C, Guerrieri-Gonzaga A, Serrano D, Cazzaniga M, Mora S, et al. Randomized double-blind 2x2 trial of low-dose tamoxifen and fenretinide for breast cancer prevention in high-risk premenopausal women. *Journal of Clinical Oncology* 2009;**27**(23):3749-56.

Ekpo 2016 *{published data only}*

Ekpo EU, Brennan PC, Mello-Thoms C, McEntee MF. Relationship between breast density and selective estrogen-receptor modulators, aromatase inhibitors, physical activity, and diet: a systematic review. *Integrative Cancer Therapies* 2016;**15**(2):127–44.

Engmann 2017 *{published data only}*

Engmann NJ, Scott CG, Jensen MR, Ma L, Brandt KR, Mahmoudzadeh AP, et al. Longitudinal changes in volumetric breast density with tamoxifen and aromatase inhibitors. *Cancer Epidemiology, Biomarkers & Prevention* 2017;**26**(6):930-7.

Eriksson 2018 *{published data only}*

Eriksson L, He W, Eriksson M, Humphreys K, Bergh J, Hall P, et al. Adjuvant therapy and mammographic density changes in women with breast cancer. *JNCI Cancer Spectrum* 2018;**2**(4):pky071.

Fabian 2006 *{published data only}*

Fabian CJ, Kimler BF. Mammographic density: use in risk assessment and as a biomarker in prevention trials. *Journal of Nutrition* 2006;**136**(10):2705S-8S.

Fabian 2007 *{published data only}*

Fabian CJ, Kimler BF. Use of biomarkers for breast cancer risk assessment and prevention. *Journal of Steroid Biochemistry and Molecular Biology* 2007;**106**(1-5):31-9.

Fabian 2016 *{published data only}*

Fabian CJ, Kimler BF. Incorporating biomarkers in studies of chemoprevention. *Advances in Experimental Medicine and Biology* 2016;**882**:69-94.

Ghosh 2010 *{published data only}*

Ghosh K, Vachon CM. Mammographic breast density, endocrine therapies, and breast cancer risk. *Sexuality, Reproduction and Menopause* 2010;**8**(1):34-9.

Kanbayti 2019 *{published data only}*

Kanbayti IH, Rae WI, McEntee MF, Ekpo EU. Are mammographic density phenotypes associated with breast cancer treatment response and clinical outcomes? A systematic review and meta-analysis. *Breast* 2019;**47**:62-76.

Kim 2014 *{published data only}*

Kim JY, Cho N, Jeyanth JX, Kim WH, Lee SH, Gweon HM, et al. Smaller reduction in 3D breast density associated with subsequent cancer recurrence in patients with breast cancer receiving adjuvant tamoxifen therapy. *American Journal of Roentgenology* 2014;**202**(4):912-21.

Kmietowicz 2013 *{published data only}*

Kmietowicz Z. Breast density on mammography can help assess response to tamoxifen and predict survival, finds study. *BMJ* 2013;**346**:f2645.

Macis 2011 *{published data only}*

Macis D, Gandini S, Guerrieri-Gonzaga A, Harriet J, Serrano D, Cazzaniga M, et al. Associations between a polymorphic locus in the promoter region of the IGFBP-3 gene, circulating IGFs, mammographic breast density and survival in a 2x2 trial of low-dose tamoxifen and fenretinide in premenopausal women. In: 102nd Annual Meeting of the American Association for Cancer Research; 2011 Apr 2-6; Orlando (FL). Vol. 71 (8 Suppl). Philadelphia: AACR; Cancer Research, 2011:Abstract number 3671.

Martin 2009 *{published data only}*

Martin L, Minkin S, Boyd N. Hormone therapy, mammographic density, and breast cancer risk. *Maturitas* 2009;**64**(1):20-6.

Martin 2016 *{published data only}*

Martin HL, Yap F, Chung K, Stone J, Redfern A. Mammographic breast-density change as a predictor of outcome in hormone receptor positive breast cancer. *Asia-Pacific Journal of Clinical Oncology* 2016;**12**(Suppl 5):72.

Mullooly 2016 *{published data only}*

Mullooly M, Pfeiffer R, Nyante S, Heckman-Stoddard BM, Perloff M, Jatoi I, et al. Mammographic density as a biosensor of tamoxifen effectiveness in adjuvant endocrine treatment of breast cancer: opportunities and implications. *Journal of Clinical Oncology* 2016;**34**(18):2093–7.

Mullooly 2018 *{published data only}*

Mullooly M, Gierach GL. The potential for mammographic breast density change as a biosensor of adjuvant tamoxifen therapy adherence and response. *JNCI Cancer Spectrum* 2018;**2**(4):pky072.

Mullooly 2019 *{published data only}*

Mullooly M, Nyante SJ, Pfeiffer RM, Cora R, Butcher D, Sternberg L, et al. Involution of breast lobules, mammographic breast density and prognosis among tamoxifen-treated estrogen receptor-positive breast cancer patients. *Journal of Clinical Medicine* 2019;**8**:1868.

NCT00066586 *{published data only}*

NCT00066586. Exemestane in reducing breast density in postmenopausal women at risk for breast cancer. clinicaltrials.gov/ct2/show/NCT00066586 (first received 7 August 2003).

NCT00086749 *{published data only}*

NCT00086749. Effect of tamoxifen on breast density in premenopausal women with breast cancer or high risk for breast cancer. clinicaltrials.gov/ct2/show/NCT00086749 (first received 12 July 2004).

NCT00114270 *{published data only}*

NCT00114270. Letrozole with or without zoledronate in treating healthy postmenopausal women with high breast density. clinicaltrials.gov/ct2/show/NCT00114270 (first received 14 June 2005).

NCT00238316 *{published data only}*

NCT00238316. Letrozole in preventing breast cancer in postmenopausal women who are at increased risk for breast cancer due to high breast density. clinicaltrials.gov/ct2/show/NCT00238316 (first received 13 October 2005).

NCT00316836 *{published data only}*

NCT00316836. Breast density, hormone levels, and anticancer drug levels in women with invasive breast cancer who are receiving exemestane or anastrozole. clinicaltrials.gov/ct2/show/NCT00316836 (first received 21 April 2006).

NCT00445445 *{published data only}*

NCT00445445. Changes in breast density and breast cancer risk in women with breast cancer and in healthy women. clinicaltrials.gov/ct2/show/NCT00445445 (first received 9 March 2007).

NCT00516698 *{published data only}*

NCT00516698. Changes in breast density and blood hormone levels in postmenopausal women receiving anastrozole or exemestane for breast cancer. clinicaltrials.gov/ct2/show/NCT00516698 (first received 15 August 2007).

NCT01765049 *{published data only}*

NCT01765049. Breast density change predicting response to adjuvant aromatase inhibitor (DEAR). clinicaltrials.gov/ct2/show/NCT01765049 (first received 10 January 2013).

NCT01773551 *{published data only}*

NCT01773551. Development of a quantitative tissue optical index of breast density for prediction of hormone therapy response. clinicaltrials.gov/ct2/show/NCT01773551 (first received 23 January 2013).

Ozhand 2013 *{published data only}*

Ozhand A, McKean-Cowdin R, Bernstein L, Ballard-Babash R, McTiernan A, Baumgartner KB. Short term reduction in mammographic density predicts survival in breast cancer. In: 104th Annual Meeting of the American Association for Cancer Research; 2013 Apr 6-10; Washington, DC. Vol. 73 (8 Suppl). Philadelphia: AACR; Cancer Research, 2013:Abstract number 2286.

Redfern 2016a *{published data only}*

Redfern A. Breast density change: toward tailoring of adjuvant endocrine therapy. *Asia-Pacific Journal of Clinical Oncology* 2016;**12**(Suppl 5):100.

Redfern 2016b *{published data only}*

Redfern AD, Martin HL, Stone J, Davidson JA, Yap F, Chung K. Mammographic breast density as a predictor of hormone receptor positive breast cancer recurrence: a single centre longitudinal analysis. In: 38th Annual CTRC-AACR San Antonio Breast Cancer Symposium; 2015 Dec 8-12; San Antonio (TX). Vol. 76 (4 Suppl). Philadelphia: AACR; Cancer Research, 2016:Abstract number PD1-06.

Shawky 2017 *{published data only}*

Shawky MS, Martin H, Hugo HJ, Lloyd T, Britt KL, Redfern A, et al. Mammographic density: a potential monitoring biomarker for adjuvant and preventative breast cancer endocrine therapies. *Oncotarget* 2017;**8**(3):5578-91.

Ursin 1996 *{published data only}*

Ursin G, Pike MC, Spicer DV, Porrath SA, Reitherman RW. Can mammographic densities predict effects of tamoxifen on the breast? *Journal of the National Cancer Institute* 1996;**88**(2):128-9.

Vachon 2013a *{published data only}*

Vachon CM, Ghosh K, Brandt KR. Mammographic density: potential as a risk factor and surrogate marker in the clinical setting. *Current Breast Cancer Reports* 2013;**5**(3):183-93.

Vachon 2013b *{published data only}*

Vachon C, Suman V, Brandt K, Kosel M, Buzdar A, Olson J, et al. Mammographic breast density response to aromatase inhibition. *Clinical Cancer Research* 2013;**19**(8):2144-53.

Whitman 2000 *{published data only}*

Whitman GJ, Stelling CB, Sunku K. Mammographic changes by endocrine therapy. *Seminars in Breast Disease* 2000;**3**(2):70-80.

Additional references

Altman 1998

Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Research and Treatment* 1998;**52**(1-3):289-303.

Altman 2001

Altman DG. Systematic reviews in health care: systematic reviews of evaluations of prognostic variables. *BMJ* 2001;**323**(7306):224-8.

Altman 2012

Altman D, McShane L, Sauerbrei W, Taube S. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLOS Medicine* 2012;**9**(5):e1001216.

Assi 2011

Assi V, Warwick J, Cuzick J, Duffy S. Clinical and epidemiological issues in mammographic density. *Nature Reviews Clinical Oncology* 2011;**9**(1):33-40.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

24

Bland 1999

Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 1999;**8**(2):135-60.

Boyd 1998

Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. *Cancer Epidemiology and Prevention Biomarkers* 1998;**7**(12):1133-44.

Boyd 2006

Boyd N, Martin L, Li Q, Sun L, Chiarelli A, Hislop G, et al. Mammographic density as a surrogate marker for the effects of hormone therapy on risk of breast cancer. *Cancer Epidemiology, Biomarkers & Prevention* 2006;**15**(5):961-6.

Byng 1994

Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Physics in Medicine and Biology* 1994;**39**(10):1629-38.

Byrne 2017

Byrne C, Ursin G, Martin CF, Peck JD, Cole EB, Zeng D, et al. Mammographic density change with estrogen and progestin therapy and breast cancer risk. *Journal of the National Cancer Institute* 2017;**109**(9):1-7.

CGHFBC 2019

Collaborative Group on Hormonal Factors in Breast Cancer (CGHFBC). Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence. *Lancet* 2019;**394**(10204):1159-68.

Covidence [Computer program]

Veritas Health Innovation Covidence. Version accessed 28 June 2018. Melbourne, Australia: Veritas Health Innovation. Available at covidence.org.

Cuzick 2004

Cuzick J, Warwick J, Pinney E, Warren R, Duffy S. Tamoxifen and breast density in women at increased risk of breast cancer. *Journal of the National Cancer Institute* 2004;**96**(8):621-8.

Cuzick 2011b

Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *Journal of Clinical Oncology* 2011;**29**(32):4273-8.

Cuzick 2013

Cuzick J, Sestak I, Bonanni B, Costantino J, Cummings S, DeCensi A, et al. Selective oestrogen receptor modulators in prevention of breast cancer: an updated meta-analysis of individual participant data. *Lancet* 2013;**381**(9880):1827-34.

Cuzick 2014

Cuzick J, Sestak I, Forbes J, Dowsett M, Knox J, Cawthorn S, et al. Anastrozole for prevention of breast cancer in high-risk postmenopausal women (IBIS-II): an international, double-blind, randomised placebo-controlled trial. *Lancet* 2014;**383**(9922):1041-8.

Cuzick 2015

Cuzick J, Sestak I, Cawthorn S, Hamed H, Holli K, Howell A, et al. Tamoxifen for prevention of breast cancer: extended long-term follow-up of the IBIS-I breast cancer prevention trial. *Lancet Oncology* 2015;**16**(1):67-75.

D'Orsi 2013

D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System. ACR, American College of Radiology, 2013.

Davies 2011

Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, et al, Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;**378**(9793):771-84.

EBCTCG 1998

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;**351**(9114):1451-67.

EBCTCG 2015

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;**386**(10001):1341-52.

Egger 1997

Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629-34.

Ferlay 2013

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. globocan.iarc.fr (accessed 1 September 2016).

Gail 1989

Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute* 1989;**81**(24):1879-86.

Google Scholar

Google. Google Scholar. scholar.google.com (accessed 5 July 2021).

Google Translate

Google. Google Translate. translate.google.com (accessed 5 July 2021).

Goss 2011

Goss PE, Ingle JN, Alés-Martínez JE, Cheung AM, Chlebowski RT, Wactawski-Wende J, et al. Exemestane for breast-cancer

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

25

prevention in postmenopausal women. *New England Journal of Medicine* 2011;**364**:2381–91.

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Version accessed 26 June 2018. Hamilton (ON): McMaster University (developed by Evidence Prime). Available at gradepro.org.

Gram 1997

Gram IT, Funkhouser E, Tabár L. The Tabár classification of mammographic parenchymal patterns. *European Journal of Radiology* 1997;**24**(2):131-6.

Harris 2007

Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology* 2007;**25**(33):5287-312.

Harvey 1997

Harvey J, Herman C, Pinkerton J. Short-term cessation of hormone replacement therapy and improvement of mammographic specificity. *Journal of the National Cancer Institute* 1997;**89**(21):1623-5.

Hayden 2006

Hayden J, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Annals of Internal Medicine* 2006;**144**(6):427-37.

Hayden 2013

Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Annals of Internal Medicine* 2013;**158**(4):280-6.

Highnam 2010

Highnam R, Brady M, Yaffe MJ, Karssemeijer N, Harvey J. Robust Breast Composition Measurement - Volpara™. In: Martí J, Oliver A, Freixenet J, Martí R, editors(s). Digital Mammography. Vol. **6136**. Berlin: Springer, 2010:342-9.

Hingorani 2013

Hingorani A, van der Windt D, Riley R, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: Stratified medicine research. *BMJ* 2013;**346**:e5793.

Huo 2014

Huo CW, Chew GL, Britt KL, Ingman WV, Henderson MA, Hopper JL, et al. Mammographic density - a review on the current understanding of its association with breast cancer. *Breast Cancer Research and Treatment* 2014;**144**(3):479-502.

Jordan 2007

Jordan C. New insights into the metabolism of tamoxifen and its role in the treatment and prevention of breast cancer. *Steroids* 2007;**72**(13):829-42.

Komm 2014

Komm B, Mirkin S. An overview of current and emerging SERMs. *Journal of Steroid Biochemistry and Molecular Biology* 2014;**143**:207-22.

Liberati 2009

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;**339**:b2700.

McCormack 2006

McCormack V, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention* 2006;**15**(6):1159-69.

McShane 2005

McShane L, Altman D, Sauerbrei W, Taube S, Gion M, Clark G. Reporting recommendations for tumor marker prognostic studies. *Journal of Clinical Oncology* 2005;**23**(36):9067-72.

McTiernan 2005

McTiernan A, Martin C, Peck J, Aragaki A, Chlebowski R, Pisano E, et al. Estrogen-plus-progestin use and mammographic density in postmenopausal women: Women's Health Initiative randomized trial. *Journal of the National Cancer Institute* 2005;**97**(18):1366-76.

PubMed

National Center for Biotechnology Information. PubMed. pubmed.ncbi.nlm.nih.gov (accessed 5 July 2021).

Review Manager 2020 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Riley 2009

Riley RD, Sauerbrei W, Altman DG. Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis and beyond. *British Journal of Cancer* 2009;**100**(8):1219-29.

Riley 2013

Riley R, Hayden J, Steyerberg E, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. *PLOS Medicine* 2013;**10**(2):e1001380.

Rutter 2001

Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in breast density associated with initiation, discontinuation, and continuing use of hormone replacement therapy. *JAMA* 2001;**285**(2):171-6.

Schünemann 2011

Schünemann HJ, Oxman AD, Vist GE, Higgins JP, Deeks JJ, Glaziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

26

5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Schünemann 2013

Schünemann H, Brozek J, Guyatt G, Oxman A, editor(s). Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach (updated October 2013). GRADE Working Group, 2013. Available from gdt.guidelinedevelopment.org/app/handbook/handbook.html.

Sterne 2016

Sterne J, Hernán M, Reeves B, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919.

Tice 2008

Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Annals of Internal Medicine* 2008;**148**(5):337-47.

Tyrer 2004

Tyrer J, Duffy S, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in Medicine* 2004;**23**(7):1111-30.

Visvanathan 2013

Visvanathan K, Hurley P, Bantug E, Brown P, Col N, Cuzick J, et al. Use of pharmacologic interventions for breast cancer risk reduction: American Society of Clinical Oncology clinical practice guideline. *Journal of Clinical Oncology* 2013;**31**(23):2942-62.

Wolfe 1976

Wolfe JN. Breast patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology* 1976;**126**(6):1130-7.

References to other published versions of this review

Atakpa 2018

Atakpa EC, Thorat MA, Cuzick J, Brentnall AR. Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review. *Cochrane Database of Systematic Reviews* 2018, Issue 8. Art. No: CD013091. [DOI: 10.1002/14651858.CD013091]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Cuzick 2011a

Study characteristics	
Methods	Nested case-control study from within a multicentre, international randomised controlled trial (First International Breast Intervention Study, IBIS-I).
	Recruitment April 1992 to March 2001, diagnosis before 1 October 2007.
	Prognostic and predictive biomarker study.
	Prevention setting.
Participants	123 cases from the UK and Finland, 942 controls from the UK.
	Age 35 to 70 years at recruitment to IBIS-I trial.
	Premenopausal and postmenopausal women.
	Approximately twice population risk of developing breast cancer.
Comparisons	Tamoxifen 20 mg daily (n = 507), placebo daily (n = 558), 5 years of treatment.
	Visually assessed per cent density by a single reader in 5% increments.
	Density reduction 10% or more vs no change at 12- to 18-month follow-up mammogram in tamoxifen arm (prognostic biomarker).
	Density reduction 10% or more vs less than 10% at 12- to 18-month follow-up mammogram in tamoxifen arm compared with placebo arm (interaction, predictive biomarker).

Cuzick 2011a *(Continued)*

Outcome: incidence of invasive breast cancer and DCIS.

Notes	ClinicalTrials.gov Identifier: NCT00002644; ISRCTN number ISRCTN91879928; EudraCT number 2005-003091-38
	Study chair: Jack Cuzick (author)
	Sponsor: QMUL
	Additional data provided for this review.
	Funding: Cancer Research UK (UK) and the National Health and Medical Research Council (Australia).

Item	Authors' judgement	Support for judgement
Study participation	Yes	The source population broadly matched our review target population; likewise the study population. The case-control study was a subset of the trial and was reported to be similar to those not included in terms of age, background risk, and demographic factors; and tumour characteristics for cases.
Study attrition	Yes	The analysis population only included women who complied with randomised treatment allocation and were followed up in the main trial. There was little loss to follow-up in the main trial. 44 of 7154 women (0.6%) withdrew and were not included in this case-control study (referenced Cuzick 2015).
Prognostic factor measurement	Unclear	The method to measure mammographic density was valid according to review criteria. However, the definition of cutpoint for density change was not made before analysis.
Outcome measurement	Yes	Breast cancer incidence was ascertained from a clinical trial database.
Study confounding	Yes	Confounding variables from trial questionnaire. Results were adjusted for age at entry, breast density at entry, history of LCIS or atypical hyperplasia, and BMI. There was no adjustment for co-interventions or HRT use, but this had no material impact on the estimate of risk reduction associated with a reduction in breast density. There was no difference in reported compliance between cases on tamoxifen with at least 10% reduction and cases on tamoxifen with less than 10% reduction.
Statistical analysis and reporting	Yes	Sufficient presentation of data to assess adequacy of analysis. Confounding variables used in the model were based on prior knowledge and biological reasoning and selected based on a step-wise procedure and statistical significance (P < 0.05). A limitation of the reported results is that no interaction effect was provided in the paper; however, this was determined from data provided by the study authors.

Kim 2012

Study characteristics

Methods	Retrospective cohort study.
	Non-randomised endocrine therapy allocation.
	Initial ER+ breast cancer diagnosis October 2003 to December 2006.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

28

Kim 2012 *(Continued)*

Median follow-up 69 months (definition of start of follow-up time not reported).

Prognostic biomarker.

Treatment setting.

Participants	Seoul National University Hospital.
	1065 women; 127 (12%) DCIS, 938 (88%) invasive at first diagnosis.
	Age 24 to 77 years.
	Premenopausal and postmenopausal women.
Comparisons	Tamoxifen (dose not stated); AI (anastrozole or letrozole, dose not stated).
	Tamoxifen 5 years (n = 657), tamoxifen 2 to 3 years + AI (for a total 5 years) (n = 41), tamoxifen 5 years + AI (unknown total time) (n = 192), AI 5 years (n = 175), at least 2 years of treatment.
	Cumulus per cent density.
	Density reduction 5% or more vs less than 5% at 8- to 20-month follow-up mammogram in tamoxifen.
	Density reduction 5% or more vs less than 5% at 8- to 20-month follow-up mammogram in AIs.
	Outcome: recurrence.
Notes	

Item	Authors' judgement	Support for judgement
Study participation	Unclear	The source population was all women with ER+ breast cancer at the study hospital without distant metastasis of bilateral breast cancer; this is close to the predefined review target population for treatment. 2 reasons for exclusion from this population for the study were 1) did not receive adjuvant endocrine treatment or were treated for less than 2 years; 2) digital mammogram images not available. 1065 of 1542 (69%) ER+ women included. No information was provided on risk or other factors between those included and excluded.
Study attrition	Unclear	No information was provided on participant dropout, loss to follow-up, or reasons for censoring.
Prognostic factor measurement	Unclear	The measure of mammographic density was broadly valid and reliable according to the review criteria, although no information was provided on whether the reader was blinded to treatment allocation.
		5% absolute percentage density change was used to define the primary prognostic factor, which differed from an earlier study (Cuzick 2011a), and was not justified a priori.
Outcome measurement	Unclear	Stated that clinical data was from the "institution's prospectively maintained web-based database", but not clear how outcomes were ascertained.
Study confounding	No	No information was provided on when and how confounding factors were measured. Adjustments (if any) for the analyses reported by treatment were unclear, including whether results were adjusted for chemotherapy. All women were required to have received endocrine therapy for at least 2 years, but it is not clear how adherence was determined, and minimum duration of endocrine therapy was 0.9 years, so there may have been some non-compli-

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

29

Kim 2012 *(Continued)*

		ance between baseline and follow-up mammogram. There was no report on how compliance affected the relationship between density change and recurrence.
Statistical analysis and reporting	No	It was unclear if subgroup analyses included women who had received tamoxifen or AIs only; for example, it was unclear if some women included in the AI subgroup were initially on tamoxifen (between baseline and follow-up mammogram).
		It was unclear when follow-up time started in the survival analysis. If follow-up time started at baseline mammogram, not from second mammogram or 2 years (since all received endocrine therapy for at least 2 years), then those with greater changes in breast density might have had a larger time between mammograms, which would bias results towards showing a prognostic effect of density change.

Knight 2018

Study characteristics	
Methods	Case-control study.
	Controls matched on follow-up time, geographical area, birth year, diagnosis year, and ethnicity.
	Non-randomised treatment allocation.
	Initial breast cancer diagnosis 1990 to 2008, recruitment to the study between 2009 and 2012.
	Follow-up from first mammogram for a mean 8 years.
	Treatment setting.
Participants	Women from the USA and Canada recruited to the WECARE study.
	224 cases and 243 controls with mammograms at baseline and follow-up.
	Mean age 46 years at mammogram before or at first diagnosis.
	Premenopausal and postmenopausal women.
	All invasive breast cancer at first diagnosis (ER-positive and negative).
Comparisons	Tamoxifen (dose not stated), no tamoxifen treatment groups.
	Cumulus per cent density.
	Density reduction 10% or more vs less than 10% at follow-up mammogram taken 6 months to 4 years after baseline mammogram.
	Outcome: incidence of a secondary primary breast cancer (e.g. in the contralateral breast).
Notes	We excluded this study from the evidence synthesis because no data were reported for the prognostic effect of breast density change, only in those with hormone-receptor positive breast cancer who had received tamoxifen (the analysis measured the effect of breast density change on outcome after adjustment for tamoxifen use). No further information provided after contact with the study authors.

Item	**Authors' judgement**	**Support for judgement**

Knight 2018 *(Continued)*

Study participation	No	WECARE study participants were diagnosed prior to age 55 years, between 1990 and 2008, with a first primary local or regional-stage invasive breast cancer. The source population is a subset of the target population of this review (which does not restrict by age < 55 years). It also included more than a quarter of women with ER-negative breast cancer at first diagnosis; our review target population was hormone-receptor positive only.
		Women who died before they could be invited to join the study were not included in the study. It is likely that there are important differences in prognosis (for contralateral breast cancer) of women included vs those excluded. This will have resulted in women with poorer prognosis in the source population being excluded from the study sample, and increases the risk of a different relationship between density reductions and outcome for women included vs those excluded from the study due to death before recruitment to study.
		Another reason for exclusion in the study sample was lack of mammograms. Comparisons conducted between women included vs those not included from the source population showed that those excluded due to lack of mammograms were slightly younger (mean age 45 years vs 46 years), but with a similar distribution of histologic type and family history.
		It was not clear why 32 of 467 women were excluded for analysis of 435 women.
Study attrition	Yes	"Cases were also diagnosed with a second primary invasive CBC at least 2 years later with no intervening cancer diagnosis, other than a non-melanoma skin cancer or cervical carcinoma in situ. UBC controls had no history of subsequent cancer diagnosis except for non-melanoma skin cancer or cervical carcinoma in situ up to their reference date, defined as follows. The reference date for cases was the CBC diagnosis date, while for controls it was defined by adding the interval between the first breast cancer and the CBC for the matched case to the date of breast cancer diagnosis for the control, thus matching on follow-up time."
Prognostic factor measurement	Unclear	An absolute 10% change in breast density was used, as previously applied and mentioned in introduction and discussion (Sandberg 2013), but it was not clear if this was a post hoc choice of cutpoint. No information was provided on whether breast density was measured blinded to treatment (tamoxifen use).
Outcome measurement	Yes	Population-based cancer registries
Study confounding	Unclear	Confounding risk factors were obtained retrospectively through a telephone survey.
		Analysis adjusted for study centre, race (non-Hispanic white vs other), age at mammogram, menopausal status at mammogram, estimated BMI at mammogram, age at first diagnosis, age at menarche, number of full-term pregnancies, first-degree family history of breast cancer, histologic type, stage, and oestrogen receptor status of first diagnosis, chemotherapy, radiation, and tamoxifen use after first diagnosis.
		There was no report or adjustment for adherence with tamoxifen use.
Statistical analysis and reporting	No	We were unable to determine the effect of density change as a prognostic or predictive biomarker in the subset of women who received tamoxifen because this was not reported, only an effect in the entire sample that was adjusted for treatment.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

31

Ko 2013

Study characteristics

Methods	Retrospective cohort study.
	Non-randomised allocation of tamoxifen.
	Initial ER+ breast cancer diagnosis January 2003 to December 2008.
	Follow-up mean 59 months (definition of the start of follow-up time not reported).
	Prognostic biomarker.
	Treatment setting.
Participants	National Cancer Centre, Goyang, South Korea.
	1066 women.
	Age 25 to 78 years.
	Premenopausal and postmenopausal women.
	13% DCIS, 87% invasive at first diagnosis.
Comparisons	Tamoxifen (all women), at least 2 years of treatment.
	BI-RADS density.
	Reduction of at least 1 category vs no reduction of at least 1 category at 10- to 34-month follow-up mammogram.
	Outcome: recurrence.
Notes	

Item	Authors' judgement	Support for judgement
Study participation	Unclear	The study source population was all women with confirmed ER+ breast cancer who had undergone curative surgery and received tamoxifen for at least 2 years; this broadly matches the review target population.
		2402 patients were identified, of whom 1336 were excluded due to lack of mammograms or bilateral or occult breast cancer. The analysis sample included 1066 women, but it was not clear why 270 of 1336 were excluded. There was no comparison of characteristics in included and excluded patients from the n = 2402 or n = 1336 women.
Study attrition	Unclear	No information on participant dropout, loss to follow-up, or reasons for censoring
Prognostic factor measurement	Unclear	No information on restriction to contralateral breast for breast density estimation or time between baseline and follow-up mammogram. No test of intra-reader reproducibility ("as the BI-RADS density classifications are standardized").
Outcome measurement	Unclear	Outcome was measured by "additional medical record review", but the reliability and accuracy of this is unclear.

Ko 2013 *(Continued)*

Study confounding	Unclear	Stated that "clinicopathologic information on 1,066 patients [was obtained] by reviewing the prospective database of our institution". It is not clear when the confounding factors used in the analysis were measured.
		There was no adjustment for chemotherapy, although the authors reported an association with mammographic density reduction.
		All women are stated to have received tamoxifen for at least 2 years; however, it is not clear how adherence was determined, and minimum duration of endocrine therapy was 10 months, so there may have been some non-compliance between baseline and follow-up mammogram. There was no report on how compliance affected the relationship between density change and recurrence.
Statistical analysis and reporting	No	It was unclear when follow-up started in the survival analysis. It should not include when women were not actually at risk, which appears to be the maximum of the time when the second mammogram was done or 2 years (since all women received tamoxifen for at least 2 years). It is likely that analysis time was from diagnosis (the last patient entered the study December 2009; all participants had at least 2 years of tamoxifen; the minimum follow-up was 26 months, and the paper was received by the journal May 2013). The authors observed that those with longer time between mammograms had greater reductions in mammographic density. Consequently, there is a risk that the findings are biased by 'immortal' time. A longer time between mammograms would count 'at risk' for those with larger mammographic density reduction, which would lead to lower biased estimate disease rates in those with greater change in breast density.

Li 2013

Study characteristics	

Methods	Retrospective cohort study.
	Non-randomised endocrine therapy allocation.
	Initial breast cancer diagnosis 1993 to 1995, follow-up until 31 December 2008.
	Follow-up from initial diagnosis (median 14.2 years).
	Prognostic biomarker.
	Treatment setting.
Participants	Sweden.
	974 women.
	Age 50 to 74 years at first diagnosis (median age 62 to 63 years).
	Postmenopausal women only.
	All invasive at first diagnosis.
Comparisons	Tamoxifen 20 mg daily (n = 231), tamoxifen 40 mg daily (n = 123), tamoxifen 20 + 40 mg daily (n = 108), tamoxifen "other" dose daily (n = 12), median 60 months of treatment.
	Fully automated area-based method measuring absolute dense area.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)   **33**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Li 2013 *(Continued)*

Relative density reduction more than 20% vs stable density (≤ 9% increase to ≤ 10% reduction) at 6- to 36-month follow-up mammogram.

Outcome: breast cancer mortality (time to death caused by breast cancer).

Notes		

Item	Authors' judgement	Support for judgement
Study participation	Unclear	Source population for the study was all postmenopausal women diagnosed with invasive breast cancer aged 50 to 74 years. This is a different source population to the review target population, but includes relevant subsets. The study source population excludes premenopausal women from the review target population, and includes some women with hormone-receptor negative disease. 3979 women were eligible, and 3345 of 3979 approached to join the study (84%) completed the study questionnaire a mean 4 months after diagnosis (SD 1.4 months), which indicates a low risk of bias in study participation of those with early breast cancer death (due to time to administer questionnaire). 527 women were excluded due to not matching the target population (such as premenopausal and not first cancer). 1844 women were excluded from the study sample (mainly due to lack of mammograms (1603 women)). A large proportion of the potential sample was thus excluded. There was no comparison between other characteristics of included/excluded participants in study sample.
Study attrition	Yes	Follow-up information from population registry
Prognostic factor measurement	Unclear	Valid and reliable density measure according to review protocol criteria. The cutpoints chosen for density change were said to be chosen "a priori", but no justification was provided for using them; they have not been reported elsewhere.
Outcome measurement	Yes	Outcome ascertained from national population registry. Clear definitions of start of follow-up and reasons for censoring were provided.
Study confounding	Yes	Adequate adjustment: time interval between baseline and follow-up mammograms, age at baseline mammogram, ever HRT use, BMI at interview, time since menopause at baseline mammogram, oestrogen receptor status, tumour size, number of metastatic nodes, grade, radiotherapy treatment, chemotherapy treatment, and change in absolute non-dense area. Analysis for tamoxifen-treated groups was additionally adjusted for length of tamoxifen treatment.
Statistical analysis and reporting	Yes	Sufficient presentation of data to assess adequacy. Conceptual framework for adjustment was to include prognostic factors and/or those associated with density.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

34

Nyante 2015

Study characteristics

Methods	Case-control study.
	Controls matched to cases on age at diagnosis, diagnosis year, and disease stage.
	Non-randomised allocation of endocrine therapy.
	Initial ER+ breast cancer diagnosis 1990 to 2008, recruitment 1 January 1991 to 31 December 2010 (also end of follow-up).
	Prognostic biomarker.
	Treatment setting.
Participants	Kaiser Permanente Northwest, USA.
	97 cases and 252 controls.
	Age 32 to 87 years at first diagnosis.
	Premenopausal and postmenopausal women.
	All invasive at first diagnosis.
Comparisons	Tamoxifen (all women), at least 1 tamoxifen prescription started within 1 year of diagnosis.
	Cumulus per cent density.
	Density reduction more than 8.7% vs less than 0.5% at 3- to 26-month follow-up mammogram.
	Outcome: breast cancer mortality (time to death caused by breast cancer).
Notes	

Item	Authors' judgement	Support for judgement
Study participation	Unclear	The source population is women enrolled in a US health plan and diagnosed with early-stage (no distant metastasis) ER-positive unilateral primary breast cancer and treated with adjuvant tamoxifen, which is similar to the review target population. 2141 women met the eligibility criteria for the source population including 134 cases of whom 97 (72%) had mammograms available. 252 controls with mammograms available were matched.
		No information was provided on the characteristics of those cases included based on mammogram availability vs those not included.
Study attrition	Yes	"Follow-up time was calculated as the time between the first tamoxifen prescription and the earliest of the following: breast cancer death, death from another cause, last tumor registry follow-up, or December 31, 2010."
Prognostic factor measurement	Yes	Valid and reliable density and density change measures according to review protocol criteria.
Outcome measurement	Yes	From population registry, clear definitions of start of follow-up and reasons for censoring
Study confounding	Yes	Several confounding measures were recorded, but not all kept in the final adjusted model (including chemotherapy, although this was considered). The

Nyante 2015 *(Continued)*

		adjusted model included tumour differentiation, duration of tamoxifen use, smoking status.
		Study used prescription records to test whether tamoxifen non-adherence (compliance) affected the relationship between density change and breast cancer death.
Statistical analysis and reporting	Yes	Overall analysis well reported and adequate. Adjustment was based on "statistical significance" (P < 0.05) in a forward step-wise variable selection procedure. This did not include selection of chemotherapy, resulting in a small risk of bias in the reported estimate. Chemotherapy affects breast density and prognosis and is expected to influence the adjusted effect size of density change, even if not selected in the step-wise procedure.

Sandberg 2013

Study characteristics	

Methods	Case-control study.
	Controls matched on age and calendar period of first breast cancer diagnosis, adjuvant therapy, and follow-up time.
	Non-randomised allocation of endocrine therapy.
	Initial breast cancer diagnosis 1976 to 2005.
	Follow-up mean 8 years.
	Prognostic biomarker.
	Treatment setting.
Participants	Sweden.
	211 cases and 211 controls.
	Age at first diagnosis (1:1 cases and controls): ≤ 45 years (n = 74), 45 to 55 years (n = 136), 55 to 65 years (n = 112), and ≥ 65 years (n = 100).
	Premenopausal and postmenopausal women.
	All invasive at first diagnosis.
Comparisons	Endocrine therapy (1:1 cases and controls, n = 174), but specific treatments not reported.
	Fully automated area-based method measuring percentage density.
	Density reduction 10% or more vs stable density (< 10% increase to < 10% reduction) at 1- to 5-year follow-up mammogram.
	Outcome: incidence of a secondary primary breast cancer (e.g. in the contralateral breast).
Notes	

Item	Authors' judgement	Support for judgement

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

36

Sandberg 2013 *(Continued)*

Study participation	Unclear	The source population was women diagnosed with first early-stage (not metastatic) unilateral breast cancer in a population-based register of all breast cancer patients diagnosed in the Stockholm-Gotland healthcare region (N > 30,000). It is a wider population than the review target population because it includes hormone-receptor negative patients, but includes a relevant subset for the review.
		211 (46%) of the 458 eligible cases were included in the analysis; exclusions were due to lack of suitable mammograms to measure density change (including 60 cases due to lack of control mammograms). It was reported that the cases excluded due to lack of eligible mammograms "did not differ from those included in the analysis in relation to age or calendar period of first diagnosis". In addition, 66 women were excluded from the analysis if their baseline per cent density was < 10% or > 90% because they could not undergo some of the defined density changes (number excluded for the subgroup of women on endocrine therapy not reported). There was no comparison between cases included and excluded in terms of other confounding factors.
Study attrition	Yes	Follow-up information was from a population registry. Controls "had survived without distant metastasis or CBC at least as long as the time between the first and second cancer for the corresponding case".
Prognostic factor measurement	Yes	Valid and reliable density measure according to review protocol criteria.
		Valid density change cutpoint: authors reported 10% absolute percentage density change, citing Cuzick 2011a as justification.
		There was some variability in time between baseline and follow-up mammograms (follow-up mammogram ranged from 1 to 5 years after first breast cancer diagnosis). This will have included women with density change over a longer period than the review target population (up to 3 years after start of endocrine therapy), but the mean time was approximately 1.5 years, SD 0.6 years, which limits the potential effect of this.
Outcome measurement	Yes	Population-based cancer registry
Study confounding	Unclear	Controls were matched to the corresponding case on the calendar period of the first breast cancer diagnosis, age at first breast cancer diagnosis, adjuvant therapy, and follow-up time; also potentially adjusted for baseline percentage density and non-dense area (unclear in the subgroup of women on endocrine therapy).
		No adjustment for chemotherapy.
		No information on adherence to endocrine therapy.
Statistical analysis and reporting	Unclear	Analysis in the subgroup of women on endocrine therapy was a secondary objective. Due to exclusions made based on baseline breast density, the number of women analysed in the endocrine therapy group is unknown; it was also unclear if adjustments were made for baseline density in the reported odds ratio (although this is unlikely to have changed the overall result).
		Follow-up time started at baseline mammogram, but mean time until first follow-up mammogram in years was reported and was similar between cases and controls.

van Nes 2015

Study characteristics

Methods	Subcohort from a multicentre randomised controlled trial (TEAM).
	Start of TEAM trial enrolment in 2001, but time period of subcohort study unknown.
	Median follow-up from entry (6 years).
	Prognostic biomarker.
	Treatment setting.
Participants	The Netherlands.
	378 women.
	Age 45 to 91 years at baseline.
	Postmenopausal women.
	All invasive at first diagnosis.
Comparisons	Exemestane 25 mg daily for 5 years (n = 197), tamoxifen 20 mg daily for 2 to 3 years followed by 3 to 2 years of exemestane (totalling 5 years) (n = 181).
	Visually assessed per cent density (Boyd categories: 0%, < 10%, 10% to 25%, 25% to 50%, 50% to 75%, and > 75%).
	Density change calculated as the difference in Boyd categories (treated as ordinal integers) between baseline and follow-up mammogram (2 years after start of endocrine therapy, range 18 to 30 months); no cutpoint.
	Outcome: recurrence (loco-regional, distance) and incidence of a secondary primary breast cancer (e.g. in the contralateral breast).
Notes	Contact with study authors resulted in additional results provided for this review.
	ClinicalTrials.gov Identifier: NCT00279448
	Funding: Pfizer, Dutch Cancer Foundation.

Item	Authors' judgement	Support for judgement
Study participation	Yes	Source population was Dutch women (postmenopausal (aged +50 years)) with hormone-receptor positive non-metastatic breast cancer, a subset of the review target population (does not include premenopausal women). The TEAM trial recruited 2754 patients from 76 hospitals (out of a total of 92 hospitals in the Netherlands). The TEAM breast density study was a subset of the TEAM trial and included 774 patients from 13 hospitals (out of the 76 hospitals in the TEAM trial). Mammograms were available for 442 of these women, and 378 were included in the study sample; most of these exclusions were due to restriction to a per-protocol analysis (women did not comply with treatment allocation). A comparison was made between those included and not included of the 774 women; this showed some evidence that those not included were less likely to be in the 'switch to exemestane' group but little difference in other prognostic factors (note that column 2 in Table IB should be labelled 'excluded' rather than 'included').

van Nes 2015 *(Continued)*

Study attrition	Yes	Women were followed up from a trial database. No information was provided on reasons for censoring or loss to follow-up. Approximately 10% did not comply with treatment allocation and were excluded from per-protocol analysis set.
Prognostic factor measurement	Yes	Valid and reliable measure of density according to review protocol. Density change was based on mammogram 18 to 30 months after start of endocrine therapy. No density change cutpoint used.
Outcome measurement	Yes	From clinical trial database
Study confounding	No	Several confounding measures were recorded, but results were unadjusted. A per-protocol analysis was performed, so all women adhered to treatment for the length of time of their follow-up.
Statistical analysis and reporting	Unclear	Insufficient presentation of data, adjustments, and results in the publication; however, after contacting the study authors for additional information, sufficient results were provided to assess the adequacy of the analysis. Follow-up time started at 2 years after treatment (to coincide with 2-year follow-up mammogram). The study used a landmark analysis based on all patients without the event of interest and still under follow-up after 2 years, therefore the survival analysis followed up women whilst they were at risk from 2 years onwards. The analysis combined women in both treatment arms (exemestane 5 years, tamoxifen 2 to 3 years followed by exemestane 3 to 2 years), so women were on either exemestane or tamoxifen between baseline and 2-year follow-up mammogram, but results were not reported by individual endocrine therapy. The analysis was on subgroups of women who were still at risk at 2 years, but data on these subgroups were not fully reported. There were no adjustments for confounders, including age, BMI, baseline density, time between mammograms, or chemotherapy, so there is a risk of bias in the reported estimate.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Andersson 2017	Density change not measured (not defined as a measure between a baseline and a follow-up mammogram)
Atkinson 1999	Wrong outcomes
Becker 2009	Review
Boisserie-Lacroix 2008	Article not in English
Boutet 2004	Article not in English
Boyd 2001	Wrong study design
Boyd 2011	Wrong study design
Chlebowski 2003	Wrong study design
Cosmacini 1993	Abstract only; not peer reviewed, insufficient information

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review) **39**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Cuzick 2012	Wrong study design
Decensi 2004	Abstract only; not peer reviewed, insufficient information
Decensi 2009	Density change not measured (not defined as a measure between a baseline and a follow-up mammogram because repeated measures ANOVA was used; additionally digital density was calibrated by adjusting for different variables at different time points, so density was not the same at baseline and follow-up mammogram)
Ekpo 2016	Wrong outcomes
Engmann 2017	Wrong outcomes
Eriksson 2018	Wrong outcomes
Fabian 2006	Wrong study design
Fabian 2007	Wrong study design
Fabian 2016	Review
Ghosh 2010	Wrong study design
Kanbayti 2019	Review
Kim 2014	Association between MRI density method and risk not validated
Kmietowicz 2013	Editorial about another paper / not primary research data
Macis 2011	Abstract only; not peer reviewed, insufficient information
Martin 2009	Wrong study design
Martin 2016	Abstract only; not peer reviewed, insufficient information
Mullooly 2016	Review
Mullooly 2018	Editorial about another paper / not primary research data
Mullooly 2019	Duplicate study (study participants were a subset of Nyante 2015: "Women in this study were part of a previously conducted case-control study designed to evaluate the prognostic significance of MD [mammographic density] change")
NCT00066586	Wrong study design
NCT00086749	Wrong outcomes
NCT00114270	Wrong study design
NCT00238316	Wrong study design
NCT00316836	No results posted
NCT00445445	Still recruiting

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

40

Study	Reason for exclusion
NCT00516698	Wrong outcomes
NCT01765049	Still recruiting
NCT01773551	Withdrawn
Ozhand 2013	Abstract only; not peer reviewed, insufficient information
Redfern 2016a	Wrong study design
Redfern 2016b	Abstract only; not peer reviewed, insufficient information
Shawky 2017	Review
Ursin 1996	Wrong outcomes
Vachon 2013a	Wrong study design
Vachon 2013b	Wrong outcomes
Whitman 2000	Review

ANOVA: analysis of variance
MRI: magnetic resonance imaging

ADDITIONAL TABLES

Table 1. Adapted QUIPS risk of bias assessment instrument for prognostic factor studies

Biases	Issues to consider for judging overall rating of risk of bias
Instructions to assess the risk of each potential bias	**These issues will guide your thinking and judgement about the overall risk of bias within each of the six domains. These issues are taken together to inform the overall judgement of potential bias for each domain.**
1. Study participation	Goal: to judge the risk of selection bias (likelihood that relationship between density reductions and outcome is different for participants and eligible non-participants)

Source of target population	The source population or population of interest is adequately described for:
	
	 • treatment:
	   * proportion with DCIS;
	   * co-interventions (chemotherapy/targeted therapy);
	   * severity of cancer at baseline (stage, % regional spread).
	 • prevention:
	   * level of risk in population, including whether some or all are BRCA1 DNA repair associated/BRCA2 DNA repair associated (*BRCA1*/*BRCA2*) mutation carriers;
	   * prior hormone replacement therapy use;
	   * co-interventions such as diet or exercise regimens, or both.

Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample are sufficient to limit potential bias.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

41

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Table 1. Adapted QUIPS risk of bias assessment instrument for prognostic factor studies *(Continued)*

Recruitment period	Period of recruitment is adequately described.
Place of recruitment	Place of recruitment (setting and geographic location) is adequately described.
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described.
Adequate study participation	There is adequate participation in the study by eligible individuals.
Baseline characteristics	The baseline study sample (i.e. individuals entering the study) is adequately described for (treatment and prevention) age, menopausal status, co-interventions; (treatment) % DCIS, disease severity; (prevention) breast cancer risk, prior hormone replacement therapy use.
Summary study participation	The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between density change and outcome.
2. Study attrition	Goal: to judge the risk of attrition bias (likelihood that relationship between density reductions and outcome are different for completing and non-completing participants)
Proportion of baseline sample available for analysis	Response rate (i.e. proportion of study sample allocated treatment who received treatment) is adequate.
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.
Reasons and potential impact of participants lost to follow-up	Reasons for loss to follow-up are provided.

Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for age at entry and co-interventions (if any), and for:
	
	• treatment:
	  * DCIS;
	  * disease severity.
	• prevention:
	  * risk of breast cancer including BRCA1 DNA repair associated/BRCA2 DNA repair associated (*BRCA1*/*BRCA2*) carriers and testing.
	
	Whether loss to follow-up or inability to retrieve mammograms, or both, was likely related to the study outcome.

Study attrition summary	There are no important differences between these characteristics in participants who completed the study and those who did not. Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (i.e. the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between density change and outcome.
3. Prognostic factor measurement	Goal: to judge the risk of measurement bias related to how mammographic density was measured (differential measurement of mammographic density related to the level of outcome)

Definition of the prognostic factor	A clear definition or description of mammographic density is provided (e.g. including the method of measurement, if subjective then who undertook it, if treatment then whether contralateral breast was assessed).
Valid and reliable measurement of prognostic factor	Method of mammographic density change measurement is adequately valid and reliable to limit misclassification bias (e.g. may include relevant outside sources of information on measurement properties; also characteristics, such as measurement blinded to case status). Continuous variables are reported or appropriate cutpoints (i.e. not data-dependent (except for percentiles)) are used.
Method and setting of prognostic factor measurement	The method and setting of measurement of mammographic density is the same for all study participants. The same mammogram type (film/digital) was used for both baseline and follow-up. The times at which baseline and follow-up mammograms were conducted have low variability between participants.
Proportion of data on prognostic factor available for analysis	An adequate proportion of the study sample has complete data for the change in mammographic density variable.
Method used for missing data	Appropriate methods of imputation are used for missing mammographic density data.
Summary	Prognostic factor is adequately measured in study participants to sufficiently limit potential bias.
4. Outcome measurement	Goal: to judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the density reductions)
Definition of the outcome	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.
Valid and reliable measurement of outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias.
Method and setting of outcome measurement	The method and setting of outcome measurement is the same for all study participants, including by age and obesity groups.
Outcome measurement summary	Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.
5. Study confounding	Goal: to judge the risk of bias due to confounding (i.e. the effect of density reductions is distorted by another factor that is related to density reductions and the outcome)
Important confounders measured	Age, BMI, or another measure of adiposity is measured.
Definition of the confounding factor	Clear definitions are provided.
Valid and reliable measurement of confounders	Measurement of all important confounders is adequately valid and reliable.
Method and setting of confounding measurement	The method and setting of confounding measurement are the same for all study participants.

Table 1.   Adapted QUIPS risk of bias assessment instrument for prognostic factor studies *(Continued)*

Method used for missing data	Appropriate methods are used if imputation is employed for missing confounder data.
Appropriate accounting for confounding	The primary analysis will have been adjusted for at least age, either through the study design and analysis, or through adjustment in the analysis only; and other prognostic factors.
Study confounding summary	Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between prognostic factor and outcome.
6. Statistical analysis and reporting	Goal: to judge risk of bias related to the statistical analysis and presentation of results
Presentation of analytical strategy, model development strategy	There is sufficient presentation of data to assess the adequacy of the analysis.
Model development strategy	The strategy for model building (i.e. inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model.
Reporting of results	The selected statistical model is adequate for the design of the study. There is no selective reporting of results.
Statistical analysis and presentation summary	The statistical analysis is appropriate for the design of the study, limiting the potential for presentation of invalid or spurious results.

BMI: body mass index
DCIS: ductal carcinoma in situ

Table 2.   ROBINS-I tool (stage 1): treatment

List of confounding domains relevant to all or most studies (prognostic factors that predict whether an individual receives a SERM/AI versus no SERM/AI)
Age
Menopausal status
Body mass index
Hormone replacement therapy
ER status
Tumour size
Nodal status
HER-2 status
List of co-interventions that could differ between intervention groups and could impact on outcome
Hormone replacement therapy
Anti-HER-2 therapy
Chemotherapy
Radiotherapy

Table 2.   ROBINS-I tool (stage 1): treatment *(Continued)*

    Mastectomy

ER: oestrogen receptor
HER: human epidermal growth factor receptor
SERM/AI: selective oestrogen receptor modulator/aromatase inhibitor

Table 3.   ROBINS-I tool (stage 1): prevention

List of confounding domains relevant to all or most studies (prognostic factors that predict whether an individual receives a SERM/AI versus no SERM/AI)
Age
Menopausal status
Body mass index
Family history of disease
Hormone replacement therapy use
Benign breast disease
Previous cancer other than breast cancer
Ethnicity

List of co-interventions that could differ between intervention groups and could impact on outcome
Hormone replacement therapy
Risk-reducing surgery

SERM/AI: selective oestrogen receptor modulator/aromatase inhibitor

A P P E N D I C E S

Appendix 1. CENTRAL

#1 MeSH descriptor: [Selective Estrogen Receptor Modulators] explode all trees
#2 MeSH descriptor: [Aromatase Inhibitors] explode all trees
#3 MeSH descriptor: [Tamoxifen] explode all trees
#4 tamoxifen
#5 MeSH descriptor: [Raloxifene Hydrochloride] explode all trees
#6 raloxifene or lasofoxifene or arzoxifene or droloxifene or bazedoxifene or fulvestrant or anastrozole or letrozole or exemestane
#7 MeSH descriptor: [Breast Neoplasms] explode all trees and with qualifier(s): [therapy - TH]
#8 ((mammogr* or breast or mammary) near dens*):ti,ab,kw
#9 MeSH descriptor: [Mammography] explode all trees
#10 MeSH descriptor: [Mammary Glands, Human] explode all trees
#11 (dens*):ti,ab,kw
#12 MeSH descriptor: [Breast Density] this term only
#13 {OR #1-#7}
#14 (#9 OR #10) AND #11
#15 #8 OR #12 OR #14
#16 #14 AND #15

Appendix 2. MEDLINE via OvidSP

1	exp Selective Estrogen Receptor Modulators/
2	exp Aromatase Inhibitors/
3	exp TAMOXIFEN/
4	tamoxifen.mp.
5	exp Raloxifene Hydrochloride/
6	raloxifene.mp.
7	lasofoxifene.mp.
8	arzoxifene.mp.
9	droloxifene.mp.
10	bazedoxifene.mp.
11	fulvestrant.mp.
12	anastrozole.mp.
13	letrozole.mp.
14	exemestane.mp.
15	breast neoplasms/th
16	or/1-15
17	exp Breast Density/
18	exp MAMMOGRAPHY/
19	exp Mammary Glands, Human/
20	((mammogr* or breast or mammary) adj6 dens*).tw.
21	dens*.tw.
22	(18 or 19) and 21
23	17 or 20 or 22
24	16 and 23
25	Animals/ not Humans/
26	24 not 25
27	limit 26 to yr="1996 -Current"

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

46

Appendix 3. Embase via OvidSP

#	Searches
1	exp selective estrogen receptor modulator/
2	exp aromatase inhibitor/
3	exp tamoxifen/
4	tamoxifen.ti,ab.
5	exp raloxifene/
6	raloxifene.ti,ab.
7	exp lasofoxifene/
8	lasofoxifene.ti,ab.
9	exp arzoxifene/
10	arzoxifene.ti,ab.
11	exp droloxifene/
12	droloxifene.ti,ab.
13	exp bazedoxifene/
14	bazedoxifene.ti,ab.
15	exp fulvestrant/
16	fulvestrant.ti,ab.
17	exp anastrozole/
18	anastrozole.ti,ab.
19	exp letrozole/
20	letrozole.ti,ab.
21	exp exemestane/
22	exemestane.ti,ab.
23	breast neoplasms/th
24	or/1-23
25	exp breast density/
26	((mammogr$ or breast or mammary) adj6 dens$).ti,ab.

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

47

(Continued)

27	dens$.ti,ab.
28	exp mammography/
29	exp mammary gland/
30	27 and (28 or 29)
31	25 or 26 or 30
32	24 and 31
33	limit 32 to (human and (conference abstracts or embase) and yr="1996 -Current")

Appendix 4. WHO ICTRP

Basic search:

1. breast density OR mammographic density

Advanced search:

Title: density

Condition: breast cancer

Intervention: selective oestrogen receptor modulator OR serm OR aromatase inhibitor OR tamoxifen OR raloxifene OR lasofoxifene OR arzoxifene OR droloxifene OR bazedoxifene OR fulvestrant OR anastrozole OR letrozole OR exemestane

Recruitment status: ALL

Appendix 5. ClinicalTrials.gov

Advanced search:

Condition or disease: breast cancer

Other terms: breast density OR mammographic density

Study type: All studies

Study results: All studies

Sex: All

Intervention/treatment: selective oestrogen receptor modulator OR serm OR aromatase inhibitor OR tamoxifen OR raloxifene OR lasofoxifene OR arzoxifene OR droloxifene OR bazedoxifene OR fulvestrant OR anastrozole OR letrozole OR exemestane

Appendix 6. Abbreviations for 'Characteristics of included studies'

AI: Aromatase inhibitor

BI-RADS: Breast Imaging Reporting and Data System

BMI: Body mass index

CBC: Contralateral breast cancer

DCIS: Ductal carcinoma in situ

ER+: Oestrogen receptor positive

HRT: Hormone replacement therapy

Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

48

LCIS: Lobular carcinoma in situ

SD: Standard deviation

UBC: Unilateral breast cancer

HISTORY

Protocol first published: Issue 8, 2018

CONTRIBUTIONS OF AUTHORS

All authors discussed and agreed upon the objectives and aims of the review. AB and EA drafted the report, which was updated to address critiques from MT, JC, and peer reviewers. All authors agreed upon the final version of the review.

DECLARATIONS OF INTEREST

We planned that if any studies eligible for inclusion in the review were co-authored by any of the authors of this Cochrane Review, these authors would not be involved in the risk of bias assessment or data extraction of these articles.

SOURCES OF SUPPORT

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the review protocol, we did not list incidence of a secondary primary breast cancer as an outcome to be included in the summary of findings tables. It was agreed that this outcome should be reported because of the limited number of overall studies and outcomes (we found no studies for four of the seven predefined outcomes to be included in the summary of findings tables). Additionally, we presented the recurrence outcome overall and separately by its location (loco-regional or distance) to reflect the precise outcomes reported in the included studies. Not all a priori subgroup analyses were reported due to lack of sufficient included studies for those subgroups.