HEALTH TECHNOLOGY ASSESSMENT

VOLUME 17 ISSUE 13 MARCH 2013
ISSN 1366-5278

Systematic review of the psychological consequences of false-positive screening mammograms

M Bond, T Pavey, K Welch, C Cooper, R Garside, S Dean and C Hyde

National Institute for Health Research

DOI 10.3310/hta17130

Systematic review of the psychological consequences of false-positive screening mammograms

M Bond,[1]* T Pavey,[1] K Welch,[2] C Cooper,[1]
R Garside,[1] S Dean[1] and C Hyde[1]

[1]Peninsula Technology Assessment Group (PenTAG), University of Exeter, Exeter, UK
[2]Karen Welch Information Consultancy, Fareham, Hampshire, UK

*Corresponding author

Declared competing interests of authors: none

Published March 2013
DOI: 10.3310/hta17130

This report should be referenced as follows:

Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, *et al*. Systematic review of the psychological consequences of false-positive screening mammograms. *Health Technol Assess* 2013;**17**(13).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.596

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (http://www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at http://www.hta.ac.uk/project/htapubs.asp. Print copies can be purchased from the individual report pages.

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Clinical Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/145/01. The contractual start date was in March 2011. The draft report began editorial review in November 2011 and was accepted for publication in June 2012. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library, produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Systematic review of the psychological consequences of false-positive screening mammograms

M Bond,[1]* T Pavey,[1] K Welch,[2] C Cooper,[1] R Garside,[1] S Dean[1] and C Hyde[1]

[1]Peninsula Technology Assessment Group (PenTAG), University of Exeter, Exeter, UK
[2]Karen Welch Information Consultancy, Fareham, Hampshire, UK

*Corresponding author

Background: In the UK, women aged 50–73 years are invited for screening by mammography every 3 years. In 2009–10, more than 2.24 million women in this age group in England were invited to take part in the programme, of whom 73% attended a screening clinic. Of these, 64,104 women were recalled for assessment. Of those recalled, 81% did not have breast cancer; these women are described as having a false-positive mammogram.

Objective: The aim of this systematic review was to identify the psychological impact on women of false-positive screening mammograms and any evidence for the effectiveness of interventions designed to reduce this impact. We were also looking for evidence of effects in subgroups of women.

Data sources: MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Health Management Information Consortium, Cochrane Central Register for Controlled Trials, Cochrane Database of Systematic Reviews, Centre for Reviews and Dissemination (CRD) Database of Abstracts of Reviews of Effects, CRD Health Technology Assessment (HTA), Cochrane Methodology, Web of Science, Science Citation Index, Social Sciences Citation Index, Conference Proceedings Citation Index-Science, Conference Proceeding Citation Index-Social Science and Humanities, PsycINFO, Cumulative Index to Nursing and Allied Health Literature, Sociological Abstracts, the International Bibliography of the Social Sciences, the British Library's Electronic Table of Contents and others. Initial searches were carried out between 8 October 2010 and 25 January 2011. Update searches were carried out on 26 October 2011 and 23 March 2012.

Review methods: Based on the inclusion criteria, titles and abstracts were screened independently by two reviewers. Retrieved papers were reviewed and selected using the same independent process. Data were extracted by one reviewer and checked by another. Each included study was assessed for risk of bias.

Results: Eleven studies were found from 4423 titles and abstracts. Studies that used disease-specific measures found a negative psychological impact lasting up to 3 years. Distress increased with the level of invasiveness of the assessment procedure. Studies using instruments designed to detect clinical levels of morbidity did not find this effect. Women with false-positive mammograms were less likely to return for the next round of screening [relative risk (RR) 0.97; 95% confidence interval (CI) 0.96 to 0.98] than those with normal mammograms, were more likely to have interval cancer [odds ratio (OR) 3.19 (95% CI 2.34 to 4.35)] and were more likely to have cancer detected at the next screening round [OR 2.15 (95% CI 1.55 to 2.98)].

Limitations: This study was limited to UK research and by the robustness of the included studies, which frequently failed to report quality indicators, for example failure to consider the risk of bias or confounding, or failure to report participants' demographic characteristics.

Conclusions: We conclude that the experience of having a false-positive screening mammogram can cause breast cancer-specific psychological distress that may endure for up to 3 years, and reduce the likelihood that women will return for their next round of mammography screening. These results should be treated cautiously owing to inherent weakness of observational designs and weaknesses in reporting. Future research should include a qualitative interview study and observational studies that compare generic and disease-specific measures, collect demographic data and include women from different social and ethnic groups.

Study registration: PROSPERO: CRD42011001345.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of abbreviations

CI	confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CNS	clinical nurse specialist
CONSORT	Consolidated Standards of Reporting Trials
CRD	Centre for Reviews and Dissemination
CWS-R	Cancer Worries Scale-Revised
FHBC	family history of breast cancer
FNA	fine-needle aspiration
GHQ	General Health Questionnaire
GP	general practitioner
HADS	Hospital Anxiety and Depression Scale
HTA	Health Technology Assessment
IBSS	International Bibliography of the Social Sciences
MeSH	medical subject heading
NHSBSP	NHS Breast Screening Programme
OPCERG	Oxford Primary Care Education Research Group
OR	odds ratio
PCQ	Psychological Consequences Questionnaire
PIMMS	Psychological Impact of Mammography Screening
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	randomised controlled trial
RR	relative risk
SD	standard deviation
STAI	State–Trait Anxiety Inventory
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
WHO	World Health Organization

All abbreviations that have been used in this report are listed here unless the abbreviation is well-known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background

Description of the health problem

In 1988, the NHS introduced a national breast screening programme (NHSBSP) for women aged 50–64 years in response to recommendations by the Forrest Committee. In 2001, the age range was expanded to 50–70 years, and currently it is being extended to 47–73 years. In the UK, women are invited for mammography every 3 years.

Rate of uptake

The Health and Social Care Information Centre's most recent statistics (2009–10) show that, in England, >2.24 million women in this age group were invited to take part in the programme, of whom 73.2% attended a screening clinic. Response rates varied according to previous screening history, with previous attenders being more likely to reattend (87.2%) than those who had received their first invitation (69.0%). Of the 1,639,953 women (aged 50–70 years) who attended screening in 2009–10 in England, 64,104 (3.9%) were recalled for further assessment. This included mammography, ultrasound, cytology, fine-needle aspiration (FNA), core biopsy and/or open biopsy of tissue. Another 1089 (0.07%) women were put on early recall and invited for screening 6 or 12 months later. Of the 64,104 women recalled, 12,525 (19.5%) were diagnosed with cancer. Thus, 51,579 women of those recalled did not have breast cancer in 2009–10 (80.5% of those recalled and 3.1% of those screened). It is this group of women who are the subject of this systematic review.

Definition of false-positive mammogram

For the purposes of this study, the definition of a false-positive mammogram is that given by the World Health Organization (WHO): 'an abnormal mammogram (one requiring further assessment) in a woman ultimately found to have no evidence of cancer'.

Objectives

The aim of this research was to identify the psychological impact on women of false-positive screening mammograms and any evidence for the effectiveness of interventions designed to reduce this impact.

The questions that this systematic review will address are:

1. What evidence is there for medium- or long-term adverse psychological consequences from false-positive screening mammograms (>1 month after assessment)?
   i. Do the types of psychological consequences differ between different groups of women?
2. What evidence is there of interventions that reduce adverse psychological consequences?

Methods

The systematic review was carried out following the principles published by the NHS Centre for Reviews and Dissemination (CRD). The study protocol can be viewed at http://www.hta.ac.uk/2510.

Inclusion and exclusion criteria

The inclusion and exclusion criteria for this systematic review are summarised as follows:

Inclusion criteria

Population (questions 1 and 2)
- Women who had received a positive result from routine mammography screening in the UK and had been invited for further assessment which showed that they did not have breast cancer.

Interventions (question 2)
- Interventions delivered to individuals to address the adverse psychological and behavioural consequences of a false-positive mammogram result.

Comparators (questions 1 and 2, respectively)
- Women who had received a negative (normal) result from routine mammography screening in the UK.
- Absence of an individual intervention in the same population.

Outcomes (questions 1 and 2)
- Psychological and behavioural outcomes and those from qualitative studies.

Setting (questions 1 and 2)
- UK.

Study design (questions 1 and 2)
- Systematic reviews, randomised, non-randomised, observational and qualitative studies.

Length of follow-up (questions 1 and 2)
At least 1 month from the 'all-clear'.

Language (questions 1 and 2)
- English language only.

Exclusion criteria
- The following types of studies were excluded: narrative reviews, editorials, opinion pieces, non-English-language papers, individual case studies and studies only reported as posters or abstracts with insufficient information to assess study quality.

Identification of studies and search strategies
The search strategy comprised the following main elements:

- electronic bibliographic databases
- internet searches
- scrutiny of references (included studies)
- contacting experts in the field.

The following electronic databases were searched in December 2010: MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Health Management Information Consortium (HMIC), Cochrane Central Register for Controlled Trials, Cochrane Database of Systematic Reviews, CRD Database of Abstracts of Reviews of Effects, CRD Health Technology Assessment (HTA), Cochrane Methodology, Web of Science, Science Citation Index, Social Sciences Citation Index, Conference Proceedings Citation Index-Science, Conference Proceeding Citation Index-Social Science and Humanities, PsycINFO, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Sociological Abstracts, the International Bibliography of the Social Sciences (IBSS) and the British Library's Electronic Table of Contents.

Ongoing trials were searched for at the UK Clinical Research Network, ControlledTrials.com, ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform and the UK Database of Uncertainties about the

Effects of Treatments. A filter was applied to capture qualitative research as well as quantitative designs. Further searches for qualitative and grey literature were run in January 2011 on the following databases: MEDLINE In-Process & Other Non-Indexed Citations, EMBASE Classic and EMBASE, British Nursing Index and Archive, Social Policy and Practice, CINAHL plus, The Cochrane Library, HMIC, PsycINFO, Applied Social Sciences Index and Abstracts, Sociological Abstracts, Web of Science, CRD and IBSS. All searches were run from inception to the search date. Bibliographies of included studies were searched for further relevant studies, including forwards and backwards chasing of citations. References were managed using Reference Manager version 11 (Thomson ResearchSoft, San Francisco, CA, USA) and EPPI-Reviewer 4 (Evidence for Policy and Practice Information and Co-ordinating Centre, University of London, London, UK). Update searches were carried out on 26 October 2011 and 23 March 2012; no new includable studies were found.

Study selection

Using the above inclusion/exclusion criteria, papers for review were selected independently by two reviewers from the titles and abstracts generated by the search strategy. Discrepancies were resolved by discussion, with the involvement of a third reviewer if necessary. Retrieved papers were reviewed and selected against the inclusion criteria by the same independent process.

Data extraction

Data regarding study design, participants, methods, outcomes, baseline characteristics and results were extracted from included studies by one reviewer using standardised data extraction forms (and checked by another reviewer). Study authors were contacted to provide missing information, as necessary.

Assessment of bias

Studies were assessed for internal and external validity according to criteria suggested by the NHS CRD Report No. 4, according to study type. Quality was evaluated using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for systematic reviews, the Consolidated Standards of Reporting Trials statement for randomised controlled trials (RCTs) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational studies.

Data analysis and synthesis

The main method of analysis was a narrative synthesis. Additional analyses were carried out using MedCalc v12 (MedCalc Software, Mariakerke, Belgium) software. The principal summary measure was relative risk (RR) with 95% confidence intervals (CIs).

Observational studies had possible sources of heterogeneity explored through assessment of the studies' populations, methods and interventions before any meta-analysis was attempted. The heterogeneity of the data did not permit meta-analysis.

Results

Number and quality of studies

We retrieved 4423 titles and abstracts, of which five systematic reviews, one meta-analysis and 11 primary studies met our inclusion criteria. None of these was about (or had subgroups of) women from different ethnic, socioeconomic or other groups within the general screening population. One study was found of women who had a false-positive mammogram and a family history of breast cancer (FHBC). One qualitative study, published only as a poster, was found.

The primary research was of variable quality, with one poor-quality RCT of an intervention and 10 observational studies. Although some studies were reasonably well reported, the majority had reporting weaknesses, including failure to report participants' demographic and other characteristics and failure to consider the possible effects of bias and confounding on the results. Indeed, in most cases, there was no

consideration of the study's methodological limitations. Therefore, the results of this systematic review must be treated with caution, because they are based on observational studies, many of which also lack methodological robustness.

Summary of results

General population

The studies of the psychological impact of false-positive mammograms in the general population gave conflicting results. When disease-specific measures were used [Psychological Consequences Questionnaire (PCQ)] an enduring negative impact was found, lasting until 35 months from the last assessment, and greatest at 5 months after the assessment (irrespective of assessment procedure). The degree of distress was related to the level of invasiveness of the assessment procedure: at 35 months, women who had a biopsy were more distressed (RR 2.07; 95% CI 1.22 to 3.52) than women who had (FNA) [RR 1.80 (95% CI 1.17 to 2.77)]; and, non-significantly, further mammography (RR 1.28; 95% CI 0.82 to 2.00). Women placed on early recall also had a greater RR of distress (RR 1.82; 95% CI 1.22 to 2.72). Conversely, when generic measures of clinical levels of general anxiety and depression were used [Hospital Anxiety and Depression Scale (HADS) and General Health Questionnaire-28], no significant differences were found between the two groups at 6 weeks after assessment and 3 months after screening.

Therefore, it may be reasonable to speculate that, for those in the general population, a false-positive mammogram may lead to breast cancer-specific psychological distress, enduring for up to 3 years, but it is unlikely that general anxiety will be detectable at clinically measurable levels.

Family history of breast cancer population

Results were slightly different for this population, with psychological distress in the false-positive group statistically significantly greater than in the normal group only at 1 month after screening (negative PCQ, difference in means 2.92; 95% CI 4.05 to 1.69). At the same time, the false-positive group also scored significantly higher on the positive PCQ than those with normal mammograms (Mann–Whitney *U*-test 51,561; $p < 0.05$). They also rated the benefits of screening more highly than those with normal mammograms at 1 month (T2) and 6 months (T3) after screening on an ad hoc questionnaire [T2: odds ratio (OR) 3.17; 95% CI 2.14 to 4.70; T3: OR 2.35; 95% CI 1.53 to 3.61]. These results may appear to be conflicting, but the summary results from the unpublished interview study suggest that the women in the false-positive group may have been rationalising their anxiety at being recalled by reassuring themselves that this meant that the programme was thorough and would detect early cancer that could be treated.

Impact of a false-positive mammogram on returning for routine screening

The evidence for the impact of having a false-positive mammogram on returning for the next screening round is conflicting. It comes mainly from four retrospective observational studies that collected data from registries and other NHS databases. The weight of evidence, in terms of the numbers of participants, suggests that women with false-positive mammograms are less likely to return for screening than women with normal mammograms. The largest study with this finding ($n = 140{,}387$) had a RR of returning of 0.97 (95% CI 0.96 to 0.98). Two studies with a combined population of 7231 found that there was no such association. Evidence from a poor-quality RCT suggests that this finding can be reversed if women are given screening invitation letters that are tailored to the outcome of their last screening (RR of returning 1.10; 95% CI 1.00 to 1.21).

Interventions to reduce the impact of false-positive mammograms

We did not find any studies that directly addressed this problem. Nevertheless, we identified two studies that investigated the information and communication needs of women who were recalled, with women wanting clear information about the reasons for recall, what their assessment would involve, and access to a breast care nurse or clinical nurse specialist (CNS) to talk through their concerns. Service satisfaction increased if women were sent a recall leaflet with their letter as participants believed that this increased their understanding of what would happen at the assessment clinic. The importance of the language used

in the recall literature was also evident with particular words and phrases reducing or increasing stress. The research by the Oxford Primary Care Education Research Group (OPCERG) was used to produce national guidelines (1998) on improving the quality of written information sent to women who are recalled for assessment.

Our results agree with those of previous systematic reviews and meta-analyses, particularly with the assertion that there can be negative psychological consequences from having a false-positive mammogram. However, we were unable to find evidence of general anxiety at clinical levels.

Additionally, it should be noted that a study by McCann *et al*. ($n = 140,387$) found that women with false-positive mammograms were at three times greater risk of interval cancer than those with normal mammograms (OR 3.19; 95% CI 2.34 to 4.35), and were more than twice as likely to have cancer detected at the next screening round (OR 2.15; 95% CI 1.55 to 2.98).

Strengths and limitations

The strengths of this systematic review are that it was conducted by an independent research team using robust methods. Comprehensive searches make it likely that we have retrieved all includable studies. Our systematic review may have been influenced by publication bias. However, there were insufficient studies in each domain to produce a meaningful funnel plot.

The robustness of the findings of this systematic review is limited by the reliability of the included studies. With the exception of one weak RCT, all the studies were observational and so subject to the risks of bias and confounding associated with these designs. This was compounded by lack of reporting key information such as the baseline characteristics.

Conclusions

We conclude that the experience of having a false-positive screening mammogram, in the general risk of breast cancer population, can cause breast cancer-specific psychological distress that may endure for up to 3 years. However, it is less likely that there will be general anxiety detectable at clinically recognisable levels. The likelihood of women experiencing distress may be determined by the degree of invasiveness of the assessment procedure, with more invasive techniques increasing the probability of psychological distress.

The strongest evidence suggests that the distress caused by a false-positive mammogram may be sufficient to deter an additional 3% of women from attending their next breast cancer screening appointment.

It is important to provide recalled women with clear, carefully worded information about the reason for the assessment and process of the assessment (but not in such detail that they become distressed without the support of the screening staff being present), and to make available a breast care nurse or CNS to talk to.

There is some evidence that having a subsequent round of screening invitation that refers to the outcome of the previous screening round may encourage women with false-positive mammograms to reattend.

For women with a FHBC, a false-positive mammogram, although increasing levels of distress, may also provide reassurance that early cancer can be detected and treated.

Research priorities

Up-to-date studies are needed that reflect current screening practice.

1. A qualitative interview study of the general population of women who have had false-positive screening mammograms, in order to understand what this experience means to them.

2. Well-designed observational studies, in the general screening population, that use disease-specific and generic outcome measures in order to determine the level of severity of negative psychological outcomes. Including studies of women from different ethnic and socioeconomic groups.
3. The routine collection of demographic information in observational studies so that future systematic reviews may be able to judge whether or not the pooling of data is possible.
4. Currently there is no standard national recall letter following a suspicious screening mammogram. There should be a national survey of the recall literature sent out from NHSBSP services to see if the national guidelines produced in 1998 are being adhered to, followed by the development of such a letter.
5. There is some evidence to suggest that there may be a relationship between tailored invitation letters for the next screening round for women who have had false-positive mammograms and reattendance. A well-designed RCT would be able to help us understand whether or not this relationship exists and a nested qualitative study would give insight into the important features of such a letter.
6. Developmental and pilot work of interventions both to relieve the distress of false-positive mammograms and to encourage women with this outcome to reattend routine screening. Promising interventions should then be tested in well-designed RCTs sufficiently powered to allow for subgroup analysis.

Study registration

This study is registered as PROSPERO: CRD42011001345.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background

Description of the health problem

In 1988, the NHS introduced a national breast screening programme (NHSBSP) for women aged 50–64 years in response to recommendations by the Forrest Committee.[1] In 2001, the age range was extended to women aged 50–70 years, currently it is being expanded to women aged 47–73 years. In the UK, women are invited for routine screening by mammography every 3 years.

Rate of uptake

The most recent statistics from the Health and Social Care Information Centre show that in 2009–10 more than 2.24 million women in this age group in England were invited to take part in the programme, of whom 73.2% attended a screening clinic.[2] The rate of response varied according to the history of previous screening. Women who had previously attended routine screening were more likely to reattend (87.2%) than those who had received their first invitation (69.0%).[2] Of the 1,639,953 women (aged 50–70 years) who attended for routine breast screening in 2009–10 in England, 64,104 (3.9%) were recalled for further assessment. This included additional mammography, ultrasound, cytology, fine-needle aspiration (FNA), core biopsy and/or open biopsy of tissue. Another 1089 women (0.07%) were put on the early recall system and invited for further screening 6 or 12 months later.[2] Of the 64,104 women recalled, 12,525 (19.5%) were diagnosed with cancer through routine screening in England in 2009–10. Thus, 51,579 women of those recalled did not have breast cancer in 2009–10 (80.5% of those recalled and 3.1% of those screened). It is this group of women who are the subject of this systematic review.

Definition of false-positive mammogram

For the purposes of this study the definition of a false-positive mammogram is that given by the World Health Organization (WHO): 'an abnormal mammogram (one requiring further assessment) in a woman ultimately found to have no evidence of cancer'.[3] This definition is rejected by some clinicians because having a positive screening mammogram is not a diagnosis of cancer in itself, but an indication that further assessment is needed. Nevertheless, for the purposes of this systematic review it is necessary to use the definition most commonly adopted in academic journals.

Incidence of breast cancer

Breast cancer is the most common cancer in the UK, with 48,034 new diagnoses in 2008.[4] It accounts for 31% of all cancers in women, with a one in nine lifetime risk.[4] The incidence of breast cancer in the separate countries within the UK can be seen in *Table 1*.

The number of cases of breast cancer in women has been steadily increasing in the UK over the last 30 years, with the annual incidence rising from 24,120 in 1978 to 47,693 in 2007. When the age of the women is standardised, the European age-standardised incidence rate increased by more than half (57%) over this time, from 77 per 100,000 in 1978 to 124 per 100,000 in 2008.[4] However, since the introduction of the national screening programme in 1988, the mortality rate has declined (*Figure 1*).

The rise in incidence has been greatest in women in higher socioeconomic groups.[5] It is thought that this may be linked to their greater use of hormone replacement therapy for menopausal symptoms[6] and the trend for having babies later in life.[7]

When the figures are broken down by age, the effects of the screening programme can be seen by the sharp increase in incidence over this time among women aged 50–64 years.[8–10] The screening programme will detect cancers that would not have been noted in the patient's lifetime (overdiagnosis) and will bring

TABLE 1 The number of new cases and rates of breast cancer in women in the UK, 2008[4]

	England	Wales	Scotland	Northern Ireland	UK
Cases (*n*)	39,681	2624	4232	1156	47,693
Crude rate per 100,000 population	151.8	171.4	158.6	127.9	152.6
Age-standardised rate (European) per 100,000 population (95% CI)	123.8 (122.6 to 125.0)	128.4 (123.5 to 133.4)	123.6 (119.8 to 127.3)	116.6 (109.9 to 123.3)	123.9 (122.8 to 125.0)

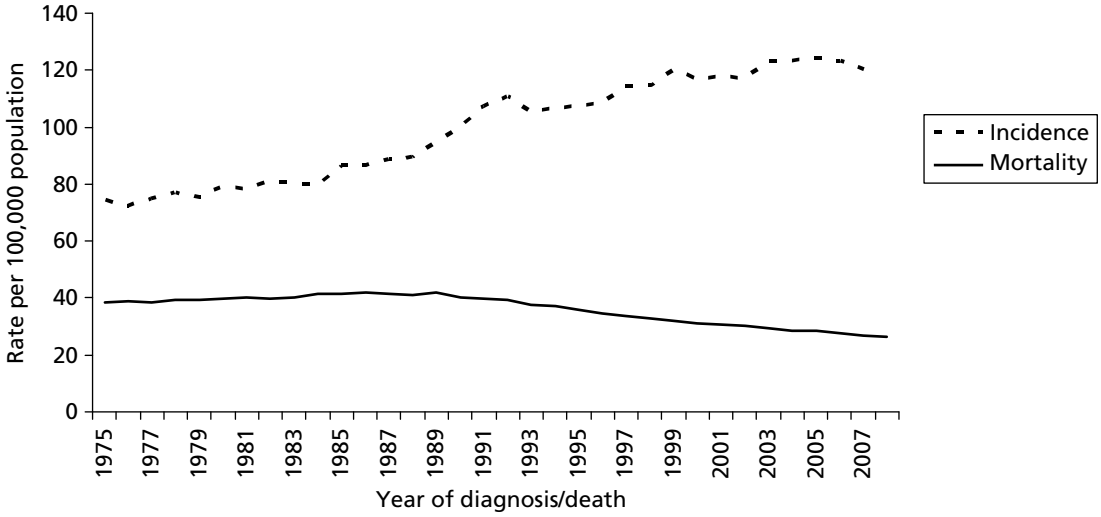

FIGURE 1 Age-standardised European incidence and mortality rates for breast cancer in women, UK, 1975–2008.[4]

forward the date of identification of cancer, finding it at an earlier stage, thus producing lead-time bias (*Figure 2*).

Mortality from breast cancer

The magnitude of the effect of mammography screening on breast cancer mortality is highly contested. Statistics for mortality from breast cancer in England are not available for 2009–10, but in 2008 a little over 10,000 women died from this disease, a rate of 26 per 100,000.[11] It can be shown that mortality from breast cancer began to decline in England at about the same time that the national breast screening programme was introduced, from about 40 per 100,000 to about 26 per 100,000 in 2008.[11] However, the exact effect that breast screening has had on breast cancer mortality is difficult to determine because it is hard to disaggregate the effects of improved treatments and other factors from the effects of screening.

However, a recent retrospective trend analysis by Autier *et al*.[12] has attempted to do this. Autier *et al*.[12] compared breast cancer mortality trends in three pairs of neighbouring European countries using WHO data from 1989 to 2006. The pairs of countries had similar demographics, quality and availability of health care. They differed in when breast cancer screening was introduced, with one country introducing it in about 1990 and the other about 10–15 years later. Autier *et al*.[12] calculated changes in breast cancer mortality using linear regressions of age-adjusted death rates. They found that although there was a wide difference in timing of the introduction of breast cancer screening in the pairs of countries there was a striking similarity in the rate of reduction of breast cancer mortality from 1990. They concluded that this reduction in breast cancer mortality was unlikely to have been the result of mammography screening. Similar findings have been replicated by the US Preventative Services Task Force,[13] who have revised their

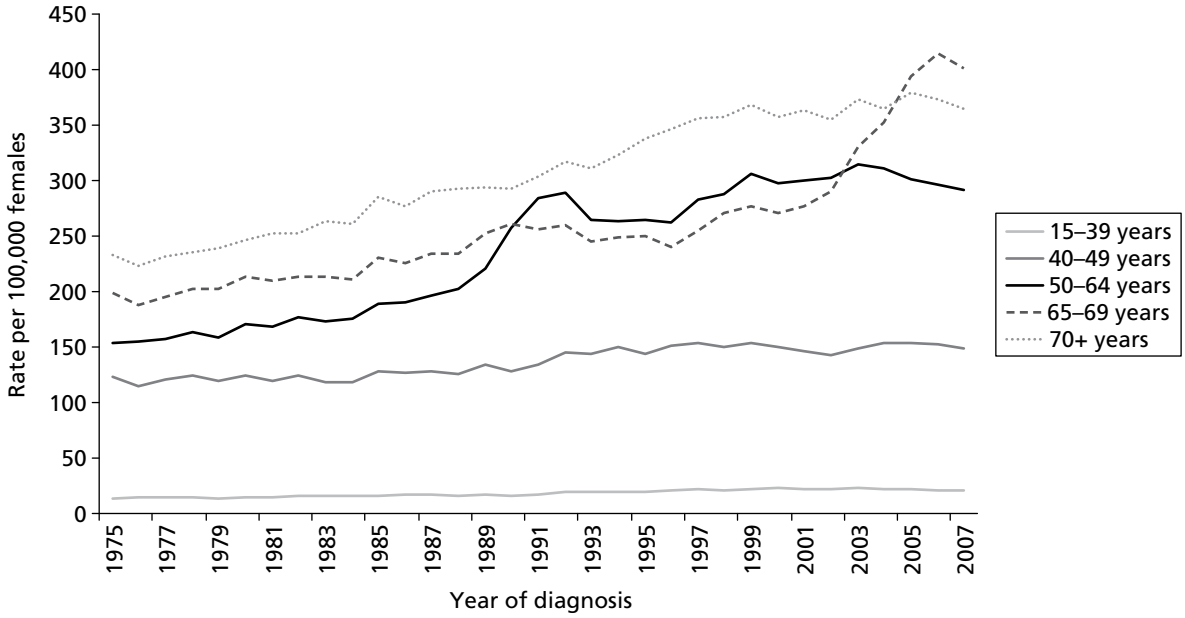

FIGURE 2 Age-specific incidence rates, female breast cancer, UK, 1975–2007.[4]

endorsement of routine screening for women aged < 50 years.[13] In a letter to the *British Medical Journal* (BMJ), Bleyer[14] presented a graph comparing US data with data from Autier *et al*.[12] Bleyer[14] concluded that improved treatment, rather than screening, is the main reason for the reduction in mortality.

This research by Autier *et al*.[12] has been criticised by de Koning[15] for being based on geographical comparisons which are unreliable and for the use of standardised all-age mortality. Further criticisms are that Autier *et al*.[12] have not accounted for the delay in time from the introduction of screening to realising its benefits.

Furthermore, in a recent systematic review of breast screening randomised controlled trials (RCTs), Gotzsche and Nielsen[16] ($n = 600{,}000$) estimated that breast screening led to a 15% reduction in breast cancer mortality, but, conversely, that there was also a 30% increase in overtreatment of women whose cancer would never become apparent in their lifetime.[16] They estimated that, over 10 years, one woman would be saved from death by breast cancer for every 2000 women invited for screening. Additionally, 10 healthy women who would have remained undiagnosed would have been treated unnecessarily for breast cancer. On top of this, for the same cohort, at least another 200 women would go through the possible distress of a false-positive outcome.[16]

Of the eight eligible trials in the Gotzsche and Nielsen review[16] (New York 1963,[17–19] Malmo 1976[20–22] and Malmo II 1978,[23] Two–County 1977,[24–26] Edinburgh 1978,[27–29] Canada 1980,[30–33] Stockholm 1981,[34–37] Goteborg 1982[38,39] and UK Age Trial 1991[40–44]), one was excluded from meta-analysis because the randomisation was seriously flawed and the data held to be unreliable (Edinburgh 1978[28,29]). Gotzsche and Nielsen[16] found that only three of the remaining trials had adequate randomisation. Pooling the data from these trials revealed no statistically significant benefit from screening on breast cancer deaths after 7 years {relative risk (RR) 0.93 [95% confidence interval (CI) 0.79 to 1.09]} or after 10 years (RR 0.90; 95% CI 0.79 to 1.02). When the data from these trials were combined with data from the other four suboptimally randomised trials, a statistically significant reduction in death from breast cancer was found after both 7 and 13 years [RR 0.81 (95% CI 0.72 to 0.90) and RR 0.81 (95% CI 0.74 to 0.87), respectively]. The pooled data from the adequately randomised trials similarly showed that there was no significant effect from breast screening on all-cause mortality after 7 years (RR 0.98; 95% CI 0.94 to 1.03) and after 13 years

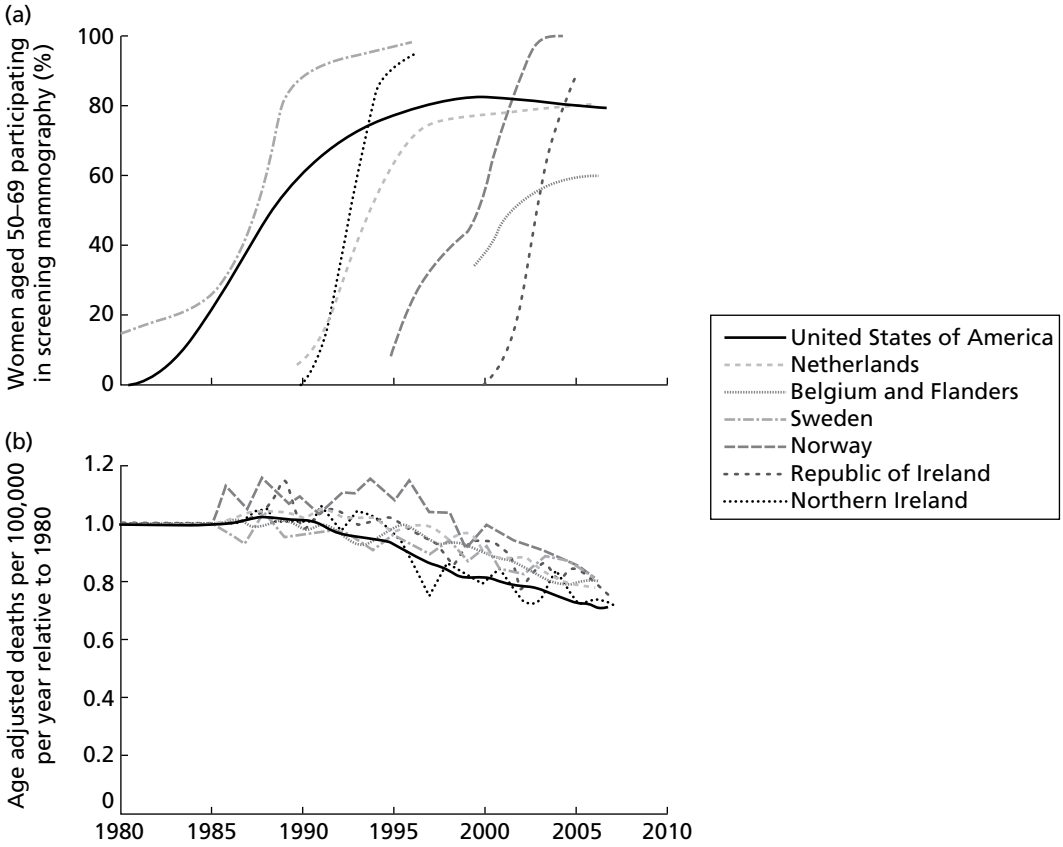

FIGURE 3 (a) Percentages of women taking part in mammography screening in each country. (b) Change in national breast cancer mortality rate relative to country's mean rate during 1980–5. Reproduced from Bleyer A. Breast cancer mortality is consistent with European data. *BMJ* 2011;**343**:d5630,[14] with permission from BMJ Publishing Group Ltd.

(RR 0.99; 95% CI 0.93 to 1.06).[16] Gotzsche and Nielsen[16] did not present data on all-cause mortality from all the included trials because the estimates were unreliable.

These results are controversial and a lively debate continues about the benefits and harms of breast cancer screening. Gotzsche and Nielsen's[16] findings have been heavily criticised by Kopans *et al.*,[45,46] who claim that the reduction in breast cancer mortality due to screening is of the order of 20–25%. Others are also highly critical and have estimated the reduction in breast cancer due to screening to be as much as 30%.[47,48] Indeed, one modelling study estimated the reduction in mortality due to screening to be between 28% and 65%.[49] The best methods for arriving at an accurate estimate of mortality reduction are also contested.[50] It is beyond the scope of this systematic review to attempt to resolve these differences. However, currently (2011), Professor Sir Mike Richards is undertaking a review to evaluate the benefits and harms of the NHS breast cancer screening services.

Significance for patients

The negative psychological impact of false-positive screening results has been documented in the fields of prenatal and cervical cancer screening.[51,52] Their impact on the psychological well-being and behaviour of women who receive false-positive results from routine mammography has been less well researched and synthesised, particularly in the UK population.

A brief examination of observational studies, looking at the psychological consequences of false-positive mammograms, showed conflicting results. Some studies indicate that, while women show increased

distress between receiving the information about the need for a follow-up appointment and receiving the all-clear, in the longer term their anxieties about breast cancer and mammography are not increased.[53–55] Other studies report that there are long-term adverse psychological consequences to receiving a false-positive mammogram.[56,57] The outcomes of studies looking at whether or not having false-positive results affects future attendance at breast screening appointments are similarly conflicting.[58–61]

A quantitative systematic review in 2007 by Brewer *et al*.[62] found that the impact of a false-positive mammogram on subsequent screening attendance varied with nationality, although the reasons for this were unclear. They also reported a varying impact on long-term psychological distress, anxiety and depression, and on other behaviours such as frequency of breast self-examination. However, their review did not report the reasons for this variation in response. Furthermore, Brewer *et al*.'s review[62] found no statistically sound studies that investigated if anxiety over a false-positive mammogram directly affects whether or not women return for routine screening or increase breast self-examination. There was little evidence about the effects on quality of life or trust of health-care services and no evidence about whether or not women who felt anxious after a false-positive screening result replaced routine screening attendance with breast self-examination.

However, the significance of receiving a false-positive mammogram result may go beyond distress and other effects on behaviour. McCann *et al*.[61] conducted a retrospective cohort study of 140,387 women, aged 49–63 years, attending NHSBSP routine screening clinics. They found that, among those women who were recalled for assessment which showed that they did not have cancer, the risk of interval cancer was increased more than threefold [rate per 1000 women screened, 9.6 (95% CI 6.8 to 12.4) compared with 3.0 (95% CI 2.7 to 3.4); odds ratio (OR) 3.19 (95% CI 2.34 to 4.35)] and these women were more than twice as likely to have cancer detected at their next routine screen in 3 years' time [rate per 1000, 8.4 (95% CI 5.8 to 10.9) vs 3.9 (95% CI 3.5 to 4.3); OR 2.15 (95% CI 1.55 to 2.98)]. This, of course, brings into question whether these women had false-positive or true-positive screening mammograms or whether or not something else explains these phenomena. It is beyond the scope of this systematic review to investigate this further.

Current related guidance

The following guidelines relate to this systematic review.

2011

NHS Breast Screening Programme 59: *Quality Assurance Guidelines for Breast Cancer Screening Radiology*.[63] These guidelines aim to raise radiology standards in breast cancer screening and relate to the transition to full-field digital mammography and the extension of the invitation to screening to women aged 47–73 years in England. They include minimising the numbers of women who are recalled and, therefore, the numbers of false-positives.

2010

NHS Breast Screening Programme 49: *Clinical Guidelines for Breast Cancer Screening Assessment*.[64] These guidelines set out minimum standards for breast cancer screening assessment. The guidelines state that this should be done using mammography or ultrasound, with clinical examination and image-guided biopsy if necessary. Women who are not diagnosed with cancer should receive written confirmation of the outcome.

2009

NHS Breast Screening Programme: *Quality Assurance Guidelines for Surgeons in Breast Cancer Screening*.[65] Among other guidelines for surgeons in breast cancer screening, these guidelines set out waiting-time targets for non-operative biopsy results to be given in < 1 week and for the time between the decision to refer for surgical assessment and the surgery taking place to be ≤ 1 week. They also aim to minimise the

numbers of benign diagnostic open surgical biopsies to < 15 per 10,000, prevalent screen and quantity of tissue taken to ≤ 20 g.

Current service provision

The UK NHSBSP is extending its service to invite women aged 47–73 years to attend for screening by mammography every 3 years. The purpose is to detect breast cancer in the general population. Contact details of eligible women are obtained through lists of registered patients by general practitioner (GP) surgery.

Mammography involves having an X-ray taken of the breasts. In the UK, two views are taken: craniocaudal (head-to-foot) and mediolateral oblique (angled side view). The mammogram is then read by two radiologists. Methods of resolving differences in opinion vary from unit to unit, but most commonly arbitration is used and a third radiologist will review the mammogram. If it is found to be normal, then the woman is put on routine recall and will receive another screening invitation in 3 years' time. There may be technical problems with the quality of the film, in which case the woman will be recalled to have the technically inadequate views repeated. Alternatively, the mammogram may show a suspicious area and the woman will be recalled for further assessment. It takes a maximum of 2 weeks from having a mammogram until a letter with the normal results is received. If further assessment is needed, the appointment for this must be within 3 weeks of the initial mammogram.

The further assessment may include another mammogram, ultrasound or core biopsy with or without FNA. FNA employs a 21-gauge needle to remove cells which are then cytologically assessed; biopsy requires a larger 14-gauge needle, which allows histologists to see the architectural context within which cells are placed and so allows more accurate diagnosis. In the UK, the lesions found are graded on a system of increasing severity: B1 is normal, B2 is benign, B3 is suspicious but probably benign, B4 is suspicious and probably cancer and B5 is cancer. B5 is further subdivided into (a) non-invasive disease, most commonly ductal carcinoma in situ (DCIS), a non-invasive, obligate precursor to invasive breast cancer, but a lesion which may not develop into invasive cancer in the lifetime of the woman, which is situated in the milk ducts; and (b) invasive cancer. It is not possible yet to assess which cases of DCIS will progress to invasive cancer and which will not do so in the lifetime of any given woman. It is likely that many such lesions would not affect the woman's lifespan. Furthermore, some lesions that are invasive may not progress to cause morbidity in the life of the individual woman. There is a debate over the amount of overdiagnosis, as this is described, that occurs.

If cancer is found, then the woman will be transferred from screening services to an oncology department. Treatment will normally be received within the NHS 62-day target from initial screening. It is usual to offer treatment for invasive cancer and non-invasive cancer as well as for many indeterminate lesions (B3 and B4).

Throughout the screening and assessment process women should have access to a clinical nurse specialist (CNS) or breast care nurse, whose role is to educate, inform and support her in a manner which is sensitive and timely and recognises her need for safety, comfort and dignity.[66] Women who are recalled are told in their letter that a breast care nurse will be available to talk to at the clinic or before the clinic by telephone if they have any concerns. Screening clinics, which may be some way from a hospital, do not routinely have breast care nurses on site. However, the invitation letter for screening will often mention the availability of a breast care nurse by telephone to allay any concerns.

Research questions

The aim of this research was to conduct a systematic review to identify the psychological impact on women of false-positive screening mammograms and any evidence for the effectiveness of interventions designed to reduce this impact. This is necessary because there is uncertainty about the nature and magnitude of their psychological impact on women, including what the predictors are of negative psychological outcomes that may affect attendance at future mammography screening. There is also a need to identify whether or not these effects differ in women from different backgrounds. This research is important because of the large number of false-positive results that come from routine mammography screening (see above).

The questions that this systematic review will address are:

1. What evidence is there for medium or long-term adverse psychological consequences from false-positive screening mammograms (>1 month after assessment)?
   i. Do the types of psychological consequences differ between different groups of women?
2. What evidence is there of interventions that reduce adverse psychological consequences?

Measurement of psychological consequences

A number of different measures of psychological morbidity are used in the primary research studies included in this systematic review. A brief summary of their characteristics is given below.

Disease specific

Psychological Consequences Questionnaire

The Psychological Consequences Questionnaire (PCQ)[67] is a reliable and validated questionnaire that was developed specifically to measure the psychological consequences of mammography screening. It comprises two subscales, one consisting of 12 statements relating to possible negative consequences of mammography screening (i.e. how often in the past week the woman has experienced loss of sleep, change of appetite, feeling depressed, being scared, feeling tense, feeling under strain, being secretive, being irritable, withdrawing socially, having difficulty in doing ordinary activities at home and work or feeling worried about the future). The second subscale relates to potential benefits from screening and has 10 items covering feeling reassured, being able to cope better with everyday life, feeling less anxious about breast cancer, feeling more hopeful and a greater sense of well-being. The statements are scored on a four-point Likert scale ranging from not at all, rarely, some of the time to quite a lot of the time, with a range of 0–36 for the negative subscale and 0–30 for the positive subscale.

Cancer Worries Scale-Revised

The Cancer Worries Scale-Revised (CWS-R)[68] is a six-item questionnaire that assesses the frequency of worries about developing cancer and how these worries affect daily mood and activities. It has been shown to be reliable and valid in at-risk populations.[68–70] The items are scored on a four-point Likert scale ranging from 1 (not at all or rarely) to 4 (almost all of the time).

Generic

Brief COPE

This is a shortened version of the COPE scale, which was developed to assess people's coping responses to stressful situations. The brief COPE[71] has been used in breast cancer patients, although the work has not been published, and in hurricane survivors. The brief version has 28 items measuring a range of coping strategies (e.g. self-distraction, active coping, denial, substance use, emotional support, instrumental

support, disengagement, venting, positive reframing, planning, humour, acceptance, religion and self-blame).

General Health Questionnaire

The General Health Questionnaire (GHQ) was designed as a screening tool for psychiatric illness in the context of general practice or general medical outpatients (i.e. non-psychiatric settings).[72] The GHQ covers the four domains of depression, anxiety, objectively observable behaviour and hypochondrias with 60 items. This instrument looks at recent experience and elicits responses on a four-point Likert scale using the statements less than usual, no more than usual, rather more than usual and much more than usual. A number of shorter versions have been developed: GHQ-12, GHQ-20, GHQ-28 and GHQ-30.[73] Items are scored 0–3 to give a total score or can be scored dichotomously with a particular threshold deemed to indicate that the respondent is 'a case'.

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) was developed to screen for psychiatric disorders in general hospital settings, excluding psychiatric wards.[74] It has two subscales, anxiety and depression, which are measured with 14 items using a four-point Likert scale (0–3). The items are totalled to give an overall score for anxiety or depression. Respondents with a score of 8–10 are considered to be 'doubtful cases' and ≥11 are considered to be 'cases'.

Life Orientation Test-Revised

The Life Orientation Test[75] is a validated measure of dispositional optimism. It has 10 items that are responded to on a five-point Likert scale, ranging from 0 (I strongly disagree) to 4 (I strongly agree).

State–Trait Anxiety Inventory

The State–Trait Anxiety Inventory (STAI) measures anxiety in adults as both a current state and an enduring trait.[76] It consists of a 20-item scale of how the respondent feels in general and a 20-item scale of how they feel now. Each item is scored on a four-point scale that ranges from 'not at all' to 'very much so'. Higher scores are related to higher anxiety.

Chapter 2 Methods

The systematic review was carried out following the principles published by the NHS Centre for Reviews and Dissemination (CRD).[77] The study protocol can be found in *Appendix 1*.

Methods for reviewing studies

Identification of studies and search strategy

The search strategy comprised the following main elements:

- searching of electronic bibliographic databases
- internet searches
- scrutiny of references of included studies
- contacting experts in the field.

The following electronic databases were searched in December 2010 for studies which met the inclusion criteria: MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Health Management Information Consortium (HMIC), Cochrane Central Register for Controlled Trials, Cochrane Database of Systematic Reviews, CRD Database of Abstracts of Reviews of Effects, CRD Health Technology Assessment (HTA), Cochrane Methodology, Web of Science, Science Citation Index, Social Sciences Citation Index, Conference Proceedings Citation Index-Science, Conference Proceeding Citation Index-Social Science and Humanities, PsycINFO, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Sociological Abstracts, the International Bibliography of the Social Sciences (IBSS) and the British Library's Electronic Table of Contents. Ongoing trials were searched for at: UK Clinical Research Network (UKCRN), ControlledTrials.com, ClinicalTrials.gov, WHO International Clinical Trials Registry Platform (ICTRP), UK Database of Uncertainties about the Effects of Treatments (DUETs), a filter was applied to capture qualitative research as well as quantitative designs. Further searches for qualitative and grey literature were run in January 2011 on the following databases: MEDLINE In-Process & Other Non-Indexed Citations, EMBASE Classic and EMBASE, British Nursing Index and Archive, Social Policy and Practice, CINAHL plus, The Cochrane Library, HMIC, PsycINFO, Applied Social Sciences Index and Abstracts, Sociological Abstracts, Web of Science, CRD and IBSS. All searches were run from inception to 25 January 2011. Bibliographies of included studies were searched for further relevant studies. References were managed using Reference Manager version 11 (Thomson ResearchSoft, San Francisco, CA, USA) and EPPI-Reviewer 4 (Evidence for Policy and Practice Information and Co-ordinating Centre, University of London, London, UK).

Initial searches were carried out between 8 October 2010 and 25 January 2011. Update searches were carried out on 26 October 2011 and 23 March 2012.

Refer to *Appendix 2* for the search strategy for MEDLINE.

Inclusion criteria

The inclusion criteria for the systematic review are summarised in *Table 2*.

Exclusion criteria

The following types of studies were excluded: narrative reviews, editorials, opinion pieces, non-English-language papers, individual case studies and studies only reported as posters or by abstract where there is insufficient information to assess the quality of the study.

TABLE 2 Inclusion criteria

Question(s)	Criteria	Specification	Notes
1 and 2	Population	Women who had received a positive result from routine mammography screening in the UK and had been invited for further assessment which showed that they did not have breast cancer	Where data permitted we looked at subgroups (including socioeconomic status and ethnic group)
2	Intervention	Those interventions delivered to individuals to address the adverse psychological and behavioural consequences of a false-positive mammogram result	These were individual interventions not group ones
1	Comparator	Women who had received a negative (normal) result from routine mammography screening in the UK	
2	Comparator	An absence of an individual intervention in the same population	
1 and 2	Outcomes	Psychological and behavioural outcomes and those from qualitative studies	Including subsequent attendance at routine mammography screening and quality of life
1 and 2	Setting	UK	Secondary care
1 and 2	Study design	Systematic reviews, randomised, non-randomised, observational and qualitative studies	We did not consider individual case studies
1 and 2	Length of follow-up	At least 1 month from the 'all-clear'	Measured over the medium- to long-term (i.e. not the immediate response to receiving a false-positive result)
1 and 2	Language	English language only	Non-English-language papers were included in the searches and screened, so that the number of potentially includable foreign-language papers is known

Study selection

Based on the above inclusion/exclusion criteria, papers were selected for review from the titles and abstracts generated by the search strategy. This was done independently by two reviewers (MB, TP); discrepancies were resolved by discussion, with the involvement of a third reviewer if necessary. Retrieved papers were again reviewed and selected against the inclusion criteria by the same independent process.

Data extraction

Data were extracted from included studies by one reviewer using standardised data extraction forms and checked by another reviewer. Authors of studies were contacted to provide missing information, as necessary. Data were gathered on the design, participants, methods, outcomes, baseline characteristics and results of the studies. The data extraction forms can be found in *Appendix 3*.

Critical appraisal – assessing risk of bias

Studies were assessed for internal and external validity according to criteria suggested by the updated NHS CRD Report No. 4, according to study type.[77,78] The quality of systematic reviews was evaluated using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[79] Individual RCTs were appraised with the Consolidated Standards of Reporting Trials (CONSORT) statement[80] and individual observational studies with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[81] There were insufficient studies in each domain to produce a meaningful assessment of publication bias with a funnel plot.

Internal validity

Consideration of internal validity addresses how well potential sources of bias and confounding are acknowledged and accounted for. Bias can be characterised as potentially undermining an experimental study in four ways: through selection bias, so that the participants in each group are dissimilar; performance bias, where the treatment of the different groups varies apart from the intervention; detection bias, which can occur if the study assessors are aware of which groups participants are in; and attrition bias, where all participants are not fully accounted for or violations of the study protocol have occurred. In particular, checks of study internal validity should address the following: whether or not there is sufficient description of the inclusion criteria, outcomes, study design, setting and the intervention to ascertain that study groups were similar in all respects and were treated in similar ways except for the intervention; if a justification for the sample size is given; if appropriate data analysis techniques were used; if dropouts and withdrawals are accounted for; if the technique used to account for missing data is described and adequate; and if assessors were blind to the group status of participants.

Another threat to validity can come from confounding. This is where an unknown agent is acting independently on the outcome being measured and the matter under investigation, so that an association appears to be occurring between the outcome measure and the matter of interest, but which is an artefact of the independent relationships.

External validity

External validity was judged according to the ability of a reader to consider the applicability of findings to a patient group and service setting. Study findings can only be generalisable if they describe a cohort that is representative of the affected population at large. Studies that appeared representative of the UK breast cancer screening population with regard to these considerations were judged to be externally valid.

Methods for analysis and synthesis

Analysis

Analysis was carried out using StatSEv12 software (TX, USA). The principal summary measure was RR with 95% CIs.

Synthesis

All study designs had a narrative synthesis. Additionally:

Randomised controlled trials and controlled trials There was only one RCT and no controlled trials.

Observational studies Observational studies had possible sources of heterogeneity carefully considered before any meta-analysis was attempted to avoid potentially spurious relationships being found. Heterogeneity was explored through assessment of the studies' populations, methods and interventions. The heterogeneity of the data did not permit meta-analysis.

Chapter 3 Results

Quantity of research available

Number and type of studies included

Electronic database searches were conducted between 8 October 2010 and 25 January 2011. The initial searches found 883 titles and abstracts after deduplication. When these were screened, 67 papers were requested for further review and two PhD theses were unobtainable. Of the 65 papers that were available, 20 were found to meet the study inclusion criteria. Four of these were systematic reviews, one was a meta-analysis and 15 were research papers; although a qualitative search filter was used, none of the papers had qualitative designs.

Further, more sensitive, qualitative searches were then conducted to see if they led to admissible studies. These searches yielded 2350 titles and abstracts after deduplication; when these had been screened 14 papers were requested. The review of these papers led to the inclusion of three primary research papers and one more systematic review. One of these papers was a summary of a nested qualitative study from an included study; this had been published only as a conference poster. Contact with the author revealed that this study had not been published in full. No published qualitative studies were found.

In order to be certain that the search strategy was picking up all includable papers, a highly sensitive further scoping search of breast cancer (and breast cancer terms) and qualitative research [as a medical subject heading (MeSH)/EMTREE/controlled syntax term] or qualitative methods (free-text terms) and qualitative terms (free text) were used. This was followed by a further scoping search of the breast cancer terms with the qualitative cluster of terms. These searches produced an enormous number of titles and abstracts ($n = 189,580$). A sample of these was screened ($n = 258$) and no further includable studies were found.

A search for grey literature produced 13 titles after deduplication. One paper was retrieved but was subsequently excluded. Breast cancer charity websites were also searched; however, no includable papers were found.

After consultation with the information specialist (CC) it was decided to do a forwards chase of citations ($n = 48$) and a backwards chase of references ($n = 50$) from one of the included systematic reviews, by Brett *et al.*,[82] to see if this was a more productive strategy for finding includable papers. This led to the retrieval of another eight papers, none of which were included. However, it was concluded that screening bibliographies and chasing citations of retrieved papers, together with contacting experts and authors, was more likely to produce includable studies than pursuing highly specific but extremely insensitive searches. This strategy is supported by Greenhalgh and Peacock,[83] who found only 30% of includable studies through searches from protocol inclusion criteria. This approach produced no further includable papers.

Update searches were carried out on 26 October 2011 and 23 March 2012; no new includable studies were found. The search strategy is available in *Appendix 2* and the updated search strategy is available from the authors.

In total, 24 papers were included (18 primary studies, five systematic reviews and one meta-analysis). The list of included papers was sent to experts in the field to confirm that there were no more relevant published papers; it was accepted that saturation had been obtained. A list of ongoing studies can be found in *Appendix 4*.

No studies were found that were either about or that had subgroups of women from different ethnic, socioeconomic or other groups. One study was found of women who had a false-positive mammogram and had a family history of breast cancer (FHBC).

A flow chart of the selection process can be found in *Figure 4*. A list of papers excluded at the paper review stage with reasons for their exclusion is available in *Appendix 5*.

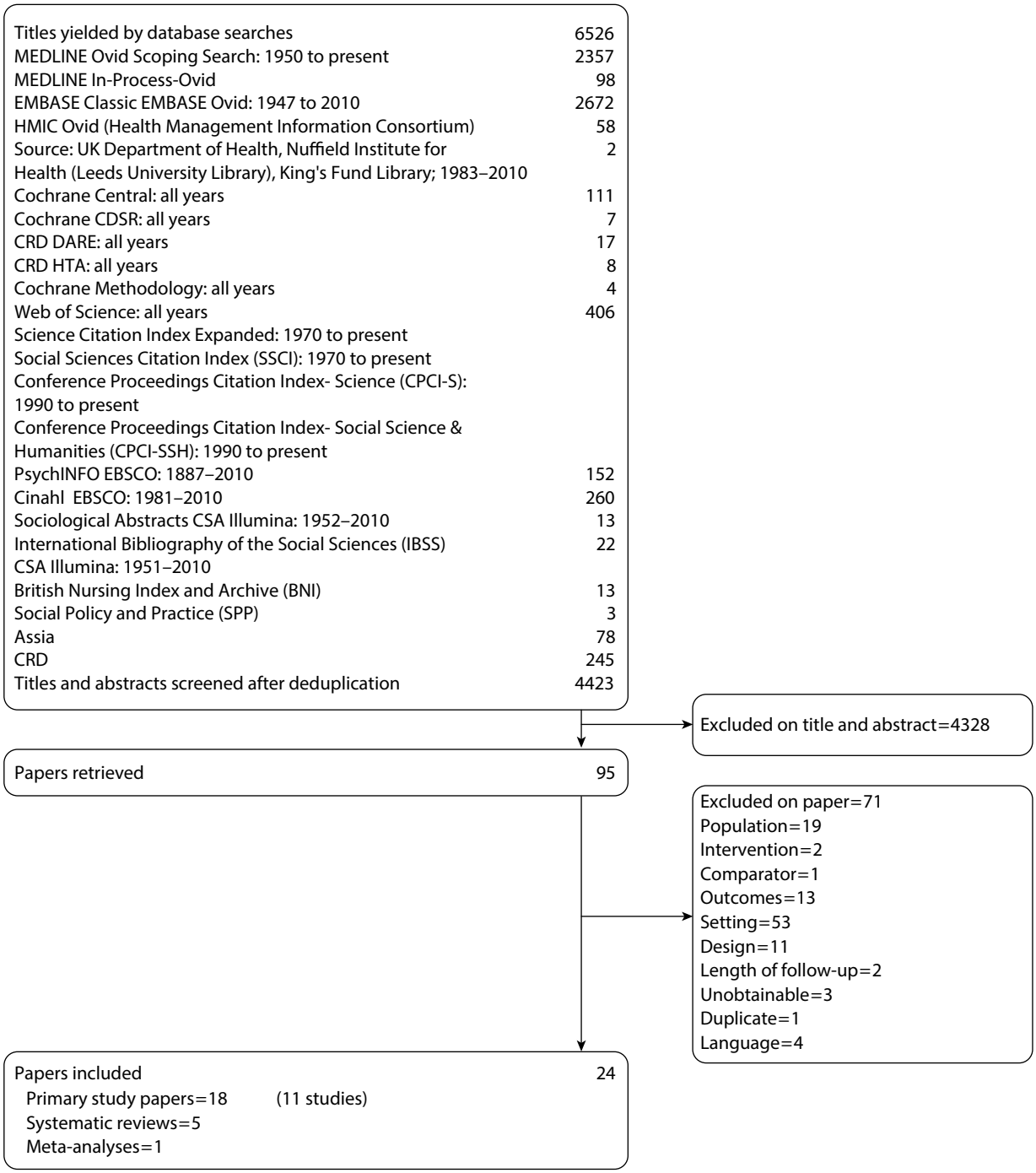

FIGURE 4 Flow chart of published evidence included in this systematic review.

Quality of studies – study characteristics and risk of bias

Systematic reviews and meta-analyses

The searches identified five systematic reviews and one meta-analysis that partly or wholly addressed the research questions of this systematic review. It was decided to evaluate the quality of all these studies against PRISMA criteria which, although not designed as a quality assessment tool, can be used for the critical appraisal of systematic reviews and meta-analyses.[79] *Table 3* provides a summary of the inclusion criteria of these studies.

Overall, the quality of methods and reporting in these systematic reviews is not high. This is despite them being conducted in the post-QUORUM (quality of reporting of meta-analyses) and, in some cases, post-PRISMA era. The exception to this is the review published by the UK HTA programme.[84] The features most commonly missing are information about access to the study protocol, presentation or access to the full electronic searches and, more worryingly, consideration of the risk of bias both within and across studies. It seems that most authors were happy to accept the results of their included studies de facto, despite being largely observational with many opportunities for bias and confounding to affect the results. A summary of the quality of the included systematic review can be found in *Table 4*.

The most recent study is a meta-analysis of the detection of the psychological impact of false-positive screening mammograms by generic and disease-specific psychosocial measures by Salz *et al*.[85] This is a secondary analysis of their previous systematic review data, published in 2007.[62] The authors were interested in comparing the difference in sensitivity of generic and disease-specific outcome measures to detect the degree of well-being experienced by women who had received a false-positive screening mammogram. Their searches found 17 observational studies of women aged ≥40 years that compared psychological outcomes between women who had had a false-positive mammogram with those whose screening results had been normal.

This meta-analysis has some weaknesses. There is no indication that any consideration of the risk of bias or confounding within or across studies was made, although the inclusion criteria and methods for study selection are clearly stated.

Additionally, Salz *et al*.[85] do not account for the number of studies screened or provide a flow chart or details of the results of the individual studies they have included. In the results section, rather than present forest plots for each analysis, they present a table of pooled effect sizes for the different criteria of outcome (e.g. distress or depression). In the discussion, although they summarise their main findings, no comment was made on the strength of the findings or their limitations at study level.

This meta-analysis was preceded by Hafslund and Nortvedt,[86] who conducted a systematic review, published in 2009, that looked at the impact on quality of life of mammography screening in women aged ≥40 years, comparing those who received a false-positive result with those whose result was normal. Psychological measures were taken in the short to medium term. The authors found 17 observational studies, which were given a narrative synthesis.

Overall, this is a poor-quality systematic review. Although methods for assessing the risk of bias in individual studies are described, insufficient information is given about the data collection process and the kind of data to be collected. The results section gives an estimation of the risk of bias within studies but the results of individual studies are inadequately presented, with little summary data. There is no assessment of the risk of bias across studies. Although the review's findings are summarised in the discussion, this is without relation to the strength of the evidence under review.

The systematic review by Brewer *et al*.[62] is of reasonable quality. Brewer *et al*.[62] were interested in the long-term effects of false-positive screening mammograms on women's attendance at their next routine breast

screening clinic. They were also interested in the impact on psychological outcomes, which they measured at least 1 month after the assessment. They found 23 observational studies that met their inclusion criteria.

The rationale and methods for conducting the study are well reported, with the notable exception of those for evaluating the risk of bias in individual studies, although methods for countering the cumulative effects of bias and assessing publication bias are given. Not surprisingly, the results section does not report the risk of bias within studies and the reporting of individual studies' results for psychological outcomes is inadequate as only a vote-counting approach is taken. However, individual, summary and pooled data for reattendance at routine screening are presented. The discussion provides a good summary of the evidence and incorporates a consideration of the limitations of the review and its components, although the strength of the evidence is not reflected on.

In the same year (2007), Armstrong *et al*.[87] published a systematic review with the aim of assessing the evidence of risks and benefits of screening mammography for women aged 40–49 years, but included a subgroup of women aged up to 71 years who had received false-positive results. This older population was compared with those with a normal screening outcome. They found 22 observational studies about false-positive mammograms.

Unfortunately, this is a poor-quality systematic review. The methods used are inadequately described with no mention of the outcomes included, methods of data synthesis or any assessment of the risk of bias across studies. The authors used a crude measure of study quality that is solely based on design, which does not consider risk of bias or confounding within studies. Furthermore, there is no indication that they have thought how bias might affect the validity of their results. The study results are particularly poorly reported; there is no account of the results of the study screening process, no results of assessment of risk of bias, inadequately described results of individual studies and a narrative synthesis almost devoid of quantified outcomes. The very brief discussion does not consider the review's limitations at any level or offer an interpretation of the results with reference to other reviews.

A systematic review by Brett *et al*.[82] was of better quality, although still lacking many of the markers of a good-quality systematic review. The authors aimed to assess the negative psychological impact of mammography screening and how long this lasted; this included the impact on women given the all-clear after screening as well as those with false-positive results. They found 52 observational studies that met their inclusion criteria.

The abstract and introduction are clear. However, the methods section is confusing to read as it contains paragraphs that should be in the introduction and results sections. The process of study selection is described, but there is no flow chart showing the progress of screening. Furthermore, the risk of bias affecting the results of individual studies or across studies is not considered in the methods or results sections. The findings of individual studies are given a comprehensive narrative summary. Recognition is given in the discussion to the shortcomings of some of the instruments used and that bias and confounding could have influenced the results. However, how this might relate to individual studies or the overall conclusions is not discussed.

The final systematic review that has been included is a good-quality HTA programme publication by Bankhead *et al*.[84] This is a broad-ranging systematic review looking at the impact of several types of screening, including mammography, on health behaviours and beliefs. Bankhead *et al*.[84] found 28 observational studies that looked at women with false-positive mammography results.

The introduction and methods are clear and comprehensive and consideration is given to the risk of bias in individual studies. However, the risk of bias across studies is not mentioned. The results section, together with the tables in the appendices, gives a full account of the findings by outcome and study, although, again, the possible effects of risk of bias across studies are not reported. The discussion section

TABLE 3 Inclusion criteria of included systematic reviews and meta-analyses

Author, year and reference	Title (no. of included studies)	Participants	Intervention	Comparator	Outcomes	Design	Exclusion criteria	Comment
Salz *et al.* 2010[85]	Meta-analysis of the effect of false-positive mammograms on generic and disease-specific psychosocial outcomes (17)	Women aged ≥40 years who had been invited to mammography screening False-positive vs normal results	Screening with mammography	No comparator	Measures of well-being and behaviour BDI, CES-D, CWS-R, GHQ, HADS, HSCL, IAS, IES, K6, PCQ, RPCS, STAI, ad hoc	Observational studies	Mammography prompted by symptoms, hypothetical studies	This study is linked to the systematic review by Brewer *et al.* 2007[62]
Hafslund and Nortvedt 2009[86]	Mammography screening from the perspective of quality of life: a review of the literature (17)	Women aged 40–74 years who had been invited to mammography screening False-positive vs normal results	Screening with mammography	No comparator	Quality of life BCAI, BDI, GHQ, HADS, HQ, PCQ, SCL-90, STAI, TTO, VAS, ad hoc	Observational studies	Women with an increased risk of breast cancer, a diagnosis of cancer, aged >74 years or <40 years, or the intervention was focused on anxiety. Non-English-language papers	Narrative synthesis
Brewer *et al.* 2007[62]	Systematic review: the long-term effects of false-positive mammograms (23)	Women aged ≥40 years who had been invited to mammography screening False-positive vs normal results	Screening with mammography	No comparator	Return for routine screening Measures of behaviour, well-being and beliefs BDI, CES-D, GHQ, HADS, HSCL, IAS, IES, K6, PCQ, SCL-90, STAI, ad hoc Measured ≥1 month after assessment	Observational studies	Mammography prompted by symptoms, hypothetical studies. Non-English-language papers	Includes a meta-analysis of the effects of false-positives on returning for routine screening

continued

TABLE 3 Inclusion criteria of included systematic reviews and meta-analyses (*continued*)

Author, year and reference	Title (no. of included studies)	Participants	Intervention	Comparator	Outcomes	Design	Exclusion criteria	Comment
Armstrong *et al.* 2007[87]	Clinical guidelines. Screening mammography in women aged 40–49 years: a systematic review for the American College of Physicians (22)	Women aged 40–49 years who had been invited to mammography screening. The subgroup of false-positive women included those > 50 years old False-positive vs normal results	Screening with mammography	No comparator	Measures of behaviour and well-being FCS, GHQ, HADS, HSCL, IES, PCQ, STAI, ad hoc	Observational studies	Case series	The false-positive studies were a subgroup of a more general review of screening mammography Narrative synthesis
Brett *et al.* 2005[82]	The psychological impact of mammographic screening. A systematic review (52)	Women who had been invited to mammography screening False-positive vs normal results	Screening with mammography	No comparator	Measures of behaviour and well-being BDI, HADS, HSCL, GHQ, PCQ, POMS, STAI, ad hoc	Observational studies	True-positive as a result of screening or symptomatic at the time of screening. Studies where the focus was on the impact of an intervention on anxiety. Non-English-language papers	Narrative synthesis
Bankhead *et al.* 2003[84]	The impact of screening on future health-promoting behaviours and health beliefs. A systematic review (28)	Women who had been invited to mammography screening False-positive vs normal results	Screening with mammography	No comparator	Health-promoting behaviours, attitudes and beliefs that result from breast screening Ad hoc	All study designs	Non-English-language papers and papers focusing on anxiety, pain or discomfort unless these affected health-promoting behaviour. Interventions to improve uptake of screening	The false-positive mammography studies were a subset of a more general review of screening that included breast, cervical and cholesterol screening Narrative synthesis

BCAI, Breast Cancer Anxiety Indicator;[88] BDI, Beck Depression Inventory;[89] CES-D, Centre for Epidemiologic Studies Depression Scale;[90] CWS-R;[68] FCS, Fear of Cancer Scale;[91] GHQ;[92] HQ, Health Questionnaire;[93] HSCL, Hopkins Symptoms Checklist;[94] IAS, Illness Attitudes Scale;[95] IES, Impact of Event Scale;[96] K6, Kessler 6;[97] POMS, Profile of Moods Scale;[98] RPCS, Rakowski Pros and Cons Scale;[99] SCL-90, Symptom-Checklist-90;[100] TTO, time trade-off; VAS, visual analogue scale.

TABLE 4 Preferred Reporting Items for Systematic Reviews and Meta-Analyses comparison of the quality of included systematic reviews

Section/topic	Item	Checklist item	Salz *et al.* 2010[85]	Hafslund and Nortvedt 2009[86]	Brewer *et al.* 2007[62]	Armstrong *et al.* 2007[87]	Brett *et al.* 2005[82]	Bankhead *et al.* 2003[84]
Title								
Title	1	Identify the report as a systematic review, meta-analysis or both	✓	✗	✓	✓	✓	✓
Abstract								
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	✓	✓	✓	✓	✓	✓
Introduction								
Rationale	3	Describe the rationale for the review in the context of what is already known	✓	✓	✓	✓	✓	✓
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes and study design	✓	✓	✓	✓	✓	✓
Methods								
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed and if available, provide registration information including registration number	✗	✗	✗	✗	✗	✗
Eligibility criteria	6	Specify study characteristics and report characteristics used as criteria for eligibility, giving rationale	✓	✓	✓	✓	✓	✓
Information sources	7	Describe all information sources in the search and date last searched	✓	✓	✓	✓	✓	✓
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	~[a]	~[a]	~[a]	~[a]	~[a]	✓

continued

TABLE 4 Preferred Reporting Items for Systematic Reviews and Meta-Analyses comparison of the quality of included systematic reviews (*continued*)

Section/topic	Item	Checklist item	Salz *et al.* 2010[85]	Hafslund and Nortvedt 2009[86]	Brewer *et al.* 2007[62]	Armstrong *et al.* 2007[87]	Brett *et al.* 2005[82]	Bankhead *et al.* 2003[84]
Study selection	9	State the process for selecting studies	✓	✓	✓	✓	✓	✓
Data collection process	10	Describe method of data extraction from reports and any processes for obtaining and confirming data from investigators	✓	✗	✓	✓	✓	✓
Data items	11	List and define all variables for which data are sort and any assumptions and simplifications made	✓	✗	✓	✗	✓	✓
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies and how this information is to be used in any data synthesis	✗	✓	✗	✗	✗	✓
Summary measures	13	State the principal summary measures	✓	✗	✓	✗	✗	✗
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measure of consistency for each meta-analysis	✓	–	✓	✗	–	–
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence	✗	✗	✓	✗	✗	✗
Additional analyses	16	Describe methods of additional analyses, if done, indicating which were pre-specified	–	–	–	–	–	–
Results								
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally from a flow diagram	✗	~[a]	✓	✗	~[c]	~[d]
Study characteristics	18	For each study, present characteristics for which data were extracted and provide the citations	✓	✓	✓	✓	✓	✓
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessments	✗	✓	✗	✗	✗	~[b]

Section/topic	Item	Checklist item	Salz *et al.* 2010[85]	Hafslund and Nortvedt 2009[86]	Brewer *et al.* 2007[62]	Armstrong *et al.* 2007[87]	Brett *et al.* 2005[82]	Bankhead *et al.* 2003[84]
Results of individual studies	20	For all outcomes considered, present for each study (a) simple summary data for each intervention group and (b) effect estimates and CIs, ideally with a forest plot	✗	~[c]	~[d]	~[f]	✓	✓
Synthesis of results	21	Present results of each meta-analysis done, including CIs and measure of consistency	✓	–	✓	–	–	–
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies	✗	✗	✓	✗	✗	✗
Additional analysis	23	Give results of additional analyses, if done	–	–	–	–	–	–
Discussion								
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome: consider their relevance for key groups	~[e]	~[h]	~[h]	~[f]	✓	✓
Limitations	25	Discuss limitation at study and outcome level and at review level	~[g]	~[h]	✓	✗	✓	✓
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implication for future research	✓	✓	✓	✗	✓	✓
Funding								
Funding	27	Describe sources of funding for the systematic review and other support and role of funders for the systematic review	✓	✓	✓	✓	✓	✓

✓, item present; ✗, item absent; ~, partially complete; – , not applicable.

a The number of hits from different databases are given, but there is no flow chart or reason for exclusions.

b Very brief comments in the results table.

c The results are inadequately presented in a largely narrative table.

d Summary data are presented in a forest plot for return to routine mammography, but other outcomes are only reported as statistically significant or not.

e Strength of evidence is not considered.

f Wholly inadequate discussion section.

g Limitations at study level are not discussed.

h Limitations at review level not considered.

gives a clear and thorough summary of the findings and the limitations of the study with suggestions for further research.

Primary research

The searches returned 11 primary research studies (18 papers) that met the inclusion criteria (including being conducted in the UK). Four of the studies were prospective cohorts,[59,101–106] one of these studies[107] included a nested interview study but this was only published as a conference poster, four were retrospective cohorts,[55,61,108,109] two had a cross-sectional design,[110–113] one of these studies[114] produced national guidelines that contained research findings and one was a RCT[115] of an intervention to improve reattendance.

Four studies looked at the psychological impact of false-positive mammograms in the normal-risk population.[55,59,103–106] One study looked at the impact of having a false-positive screening mammogram among a population of women with a FHBC.[101,102,107] Five studies looked at the impact on returning for routine mammography screening[59,61,103,108,109,115] and two studies investigated the impact of written information on distress or reattendance.[110–114]

In some of the studies, groups of women were included with characteristics outside the scope of this systematic review; in these cases only data from the study population included in this review are extracted and reported.

Although worded slightly differently, the definition of woman with a false-positive mammogram is consistent in all of the included studies and agrees with the definition used in this systematic review (i.e. a woman who is recalled for assessment of any kind on the basis of a routine screening mammogram who is not then diagnosed with breast cancer).

This section first of all gives an overview of the relationship of the studies to each other according to their domain of interest (*Figure 5*). Some papers report outcomes in more than one domain and are therefore shown accordingly in *Figure 5*. After this, the characteristics and quality of the studies are discussed according to their domain. This is followed by a summary table of the characteristics of each paper (*Table 5*). Finally, a summary in *Table 6* gives an overview of the quality of the observational studies. More detailed information about all the primary studies can be found in the data extraction forms in *Appendix 3*. A summary description of the measures used in the included primary research can be found in *Chapter 1, Measurement of psychological consequences*.

The UK research in the field of false-positive mammography screening has been dominated by the University of Oxford Primary Care Education Research Group (OPCERG) who first published a series of papers from 1997 about the information needs of women recalled for further assessment following a screening mammogram. This work then developed to investigate the psychological impact of false-positive mammograms on the general population of screened women and how this affected their reattendance at their next routine screening, and latterly looked at the psychological impact of receiving a false-positive mammogram on women who have a FHBC. Other research groups in England and Scotland have also studied these four aspects and in many cases preceded the Oxford research.

Psychological impact

The papers from the OPCERG study in this domain span 5 years, with measures being taken at 1 month (T1), 5 months (T2) and 35 months (T3) from the last screening appointment. This study is a multicentre prospective cohort at 13 NHSBSP clinics in England and Scotland. The papers follow the same cohort of women over this time but the methods used appear to vary at different time points and it is not always clear on what basis the women included in the study have been selected or how the main outcome measure, the PCQ, has been used. The study includes women with normal mammograms as the control group and compares them with women with false-positive outcomes according to the process used in

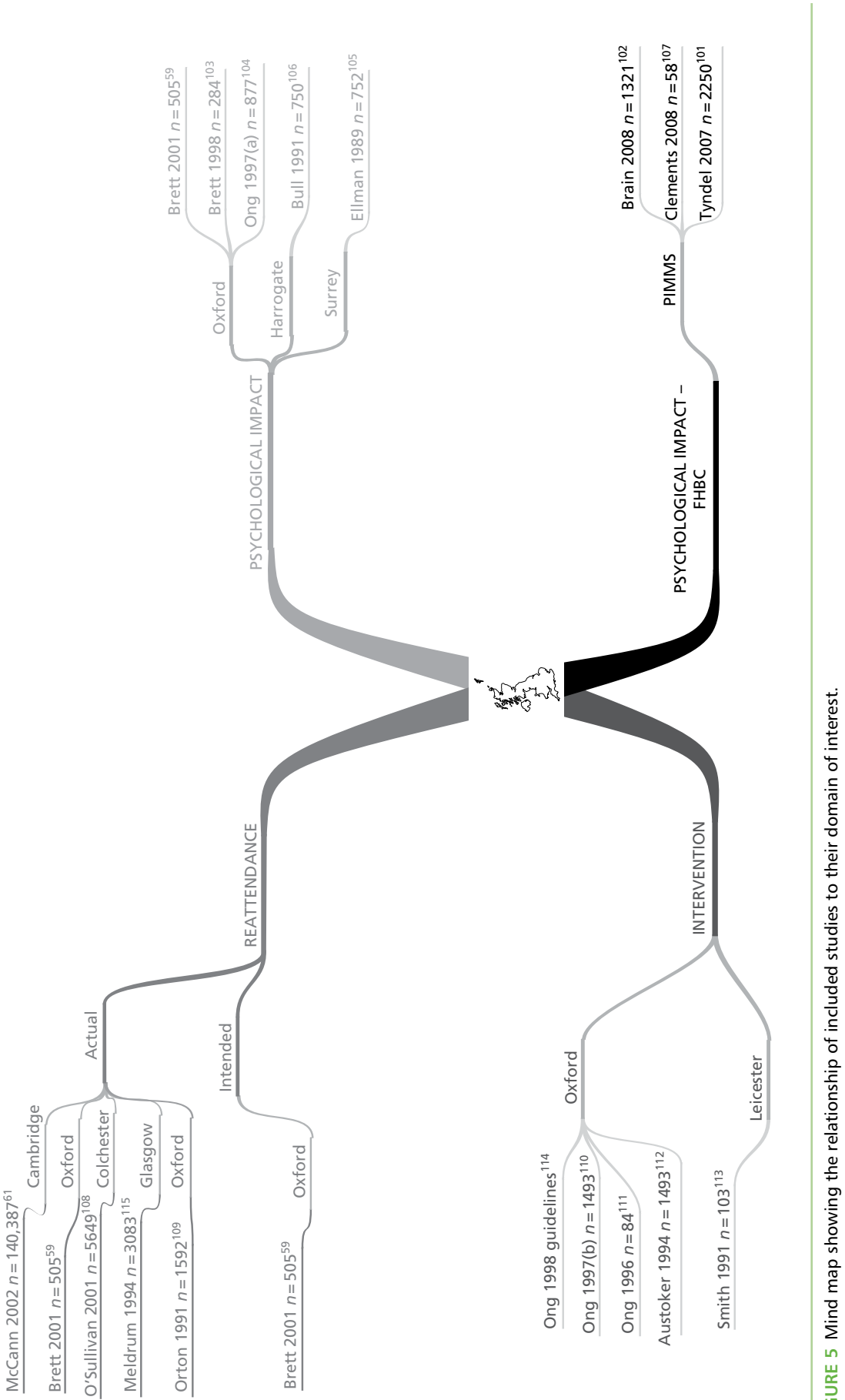

FIGURE 5 Mind map showing the relationship of included studies to their domain of interest.

TABLE 5 Summary characteristics of included primary studies by domain

Study, author, year (funding)	Design	n	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
Psychological impact									
Brett and Austoker 2001[59] (Cancer Research Campaign)	Prospective cohort, multicentre *Psychological impact*	505	Women invited for routine screening by mammography, already participating in the study at 5 months	Routine screening by mammography with a false-positive result ($n = 375$)	Routine screening by mammography with a normal result ($n = 130$)	PCQ, intention to reattend and actual reattendance satisfaction with service ad hoc questionnaire	3 years (35 months) after assessment	Aged >65 years, symptomatic referral, in another study, developed cancer	
Brett *et al.* 1998[103] (Cancer Research Campaign)	Prospective cohort, multicentre *Psychological impact*	284	Women invited for routine screening by mammography, already participating in the study at 1 month	Routine screening by mammography with a false-positive result ($n = 163$)	Routine screening by mammography with a normal result ($n = 52$)	PCQ, intention to reattend, ad hoc questionnaire	5 months after assessment	Aged >65 years, symptomatic referral, in another study, developed cancer	69 (24%) women chose not to return the questionnaire
Ong *et al.* 1997[104] (Cancer Research Campaign, NHSBSP)	Prospective cohort, multicentre *Psychological impact*	877	Women invited for routine screening by mammography who were recalled for assessment	Women placed on early recall (<3 years) ($n = 182$)	Women placed on routine recall after mammography ($n = 173$), further mammography assessment ($n = 166$), FNA ($n = 109$) or biopsy ($n = 31$)	PCQ	NA	Not reported	This study was primarily about the effects of early recall on women who had been called back for assessment after their mammogram, measures taken 1 month after assessment
Bull and Campbell 1991[106] (Yorkshire Regional Health Authority)	Prospective cohort *Psychological impact*	750	Women invited for routine screening by mammography who were recalled for assessment	Routine screening by mammography with a false-positive result ($n = 308$)	Routine screening by mammography with a normal result ($n = 420$)	Ad hoc questionnaire including frequency of breast self-examination, HADS	6 weeks after the 'all-clear'	Not reported	It is not known if the women had previously had cancer or were in a high-risk group

Study, author, year (funding)	Design	n	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
Ellman *et al.* 1989[105] (DHSS Research Management Division)	Prospective cohort *Psychological impact*	752	Women invited for routine mammography screening and those recalled for further assessment and those with symptoms being further investigated	Routine screening by mammography with a false-positive result ($n = 271$)	Routine screening by mammography with a normal result ($n = 295$), symptomatic women who did not have cancer ($n = 134$), symptomatic or recalled screened women who did have cancer ($n = 38$), history of breast cancer with or without symptoms ($n = 14$)	GHQ-28, ad hoc questionnaire	3 months after clinic attendance	Not reported	Participants also received clinical examination. Only those groups meeting the inclusion criteria will be considered in this systematic review
Psychological impact with a FHBC									
Brain *et al.* 2008[102] (Cancer Research UK)	Prospective cohort *Psychological impact*	1321	Women aged 35–49 years invited for routine annual screening by mammography with a FHBC	Routine annual screening by mammography with a false-positive result ($n = 112$)	Routine annual screening by mammography with a normal result ($n = 1174$)	Questionnaire including: CWS-R, cognitive appraisal, COPE, perceived risk of breast cancer, dispositional optimism	6 months, measures taken at 1 month before screening and 1 and 6 months after the 'all-clear'	Previous history of breast cancer or family history of ovarian cancer	This study aimed to find predictive variables of cancer-specific distress
Clements *et al.* 2008[107] (Cancer Research UK)	In-depth interviews *Psychological impact*	58	Women aged 35–49 years invited for routine annual screening by mammography with a FHBC	Routine annual screening by mammography with a false-positive result ($n = 22$)	Routine annual screening by mammography with a normal result ($n = 36$)	Reactions of women to false-positive recall and value placed on the screening programme	NA	Previous history of breast cancer or family history of ovarian cancer	It is not known if the women had previously had cancer or were in a high risk group
Tyndel *et al.* 2007[101] (Cancer Research UK)	Prospective cohort *Psychological impact*	2321	Women aged 35–49 years invited for routine annual screening by mammography with a FHBC	Routine annual screening by mammography with a false-positive result ($n = 166$)	Routine annual screening by mammography with a normal result ($n = 2084$)	CWS-R, PCQ	6 months, measures taken at 1 month before screening and 1 and 6 months after the 'all-clear'	Previous history of breast cancer or family history of ovarian cancer	

continued

TABLE 5 Summary characteristics of included primary studies by domain (*continued*)

Study, author, year (funding)	Design	n	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
Impact on reattendance									
McCann *et al.* 2002[61] (NHS Executive Eastern Region)	Retrospective cohort *Reattendance and interval cancer*	140,387	Women aged 49–63 years invited for routine breast screening by mammography	Routine screening by mammography with a false-positive result ($n = 4792$)	Routine screening by mammography with a normal result ($n = 108,617$)	Subsequent attendance at routine screening after a false-positive result and rate of interval cancer – from records	3 years	Women who were older than 63 years at follow-up	
O'Sullivan *et al.* 2001[108] (Cancer Research Campaign)	Retrospective cohort *Reattendance*	5649	Women invited for mammography screening for the second or subsequent time	Routine screening by mammography with a false-positive result ($n = 248$)	Routine screening by mammography with a normal result ($n = 5401$)	Subsequent attendance at routine screening after a false-positive result – from records	Unclear, probably from 1989 to 1997	Women invited for the first time and women who had been previously invited but had never attended	Effects of a false-positive result on reattendance for those on early recall and routine recall
Brett and Austoker 2001[59] (Cancer Research Campaign)	As above in *Psychological impact*								
Brett *et al.* 1998[103] (Cancer Research Campaign)	As above in *Psychological impact*								
Meldrum *et al.* 1994[115] (Scottish Office Home and Health Department)	RCT nested telephone interview study	3083	All women invited for second round routine mammography screening (aged 50–65 years)	Tailored invitation with a false-positive result ($n = 115$) and with a normal result ($n = 800$)	Standard invitation with a false-positive result ($n = 112$) and with a normal result ($n = 791$)	Subsequent attendance at routine screening and effect of a tailored invitation on subgroups	Not reported	Women with breast cancer and those whose screening history was not available	Trial comparing the effect of a tailored invitation on second-round screening attendance with a standard invitation

Study, author, year (funding)	Design	n	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
Orton *et al.* 1991[109] (funding not reported)	Retrospective cohort *Reattendance*	1582	Women, aged 45–64 years, invited to attend for second-round screening by mammography	Routine screening by mammography with a false-positive result ($n = 50$)	Routine screening by mammography with a normal result ($n = 1532$)	Reattendance, acceptability of screening	NA	Not reported	Data are not available for the acceptability of screening for false-positive participants
Interventions to reduce the impact of false-positive mammograms									
Ong *et al.* 1998[114] (Cancer Research Campaign, NHSBSP)	Guidelines and summary evidence from cross section, multicentre *Information intervention*	NA	Women invited for routine screening by mammography who were recalled for assessment	NA	NA	Ad hoc questionnaire and criteria for evaluating breast screening information material developed by Austoker and Ong 1994[112]	NA	Women recalled due to poor quality X-rays	National guidelines about information given prior to recall for further assessment based on the findings of Ong *et al.* 1997,[104] 1996[111] and Austoker and Ong 1994[112]
Ong and Austoker 1997[110] (Cancer Research Campaign, NHSBSP)	Cross section, multicentre *Information intervention*	1493	Women invited for routine screening by mammography who were recalled for assessment	$n = 1493$	NA	Ad hoc questionnaire	NA	Women recalled due to poor quality X-rays	Evaluation of women's experiences at the assessment clinic and their information needs there and afterwards. Discourse analysis of open questions
Ong *et al.* 1996[111] (Cancer Research Campaign, NHSBSP)	Cross section, multicentre *Information intervention*	84	UK breast screening assessment centres	Evaluation of information given in the initial letter/leaflet and prior to recall for further assessment	NA	Criteria for evaluating breast screening information material developed by Austoker and Ong 1994[112]	NA	NA	

continued

TABLE 5 Summary characteristics of included primary studies by domain (*continued*)

Study, author, year (funding)	Design	n	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
Austoker and Ong 1994[112] (Cancer Research Campaign, NHSBSP)	Cross section, multicentre *Information intervention*	1493	Women invited for routine screening by mammography who were recalled for assessment	$n = 1493$	NA	Ad hoc questionnaire	NA	Women recalled due to poor quality X-rays	Evaluation of information given prior to recall for further assessment from eight UK breast screening centres
Smith *et al*. 1991[113] (funding not reported)	Cross section *Survey*	103	Women attending assessment clinic following recall after routine mammography screening	False-positive ($N = 91$) Other outcome ($N = 12$)	NA	Ad hoc questionnaire	NA	Not reported	Survey of different versions of an invitation to return for further assessment. The responses from both groups of women are aggregated

NA, not applicable.

their further assessment (another mammogram, FNA or biopsy) or if they had been placed on early recall of 6 or 12 months following further assessment.

The first publication by OPCERG, by Ong *et al*.[104] ($n = 877$), reported primarily on the adverse psychological consequences of being placed on early recall (6 or 12 months recall rather than 3 years) at T1. These effects were measured using the PCQ negative scale which has 12 items on a four-point Likert scale (scored 0–3).[67] Participants were considered to be 'cases' if they responded positively to at least one item.

The study by Ong *et al*.[104] is a moderately good-quality study that used appropriate methods to address its aims. The authors reported almost all the key criteria specified by the STROBE statement.[81] However, they omitted to provide demographic data for all participants, not just those on early recall, which makes it difficult to interpret the results. They also failed to fully discuss the limitations of their study and its generalisability to other situations.

Following on from the study by Ong *et al*.,[104] Brett *et al*.[103] ($n = 284$) recruited from the same pool of women to find out what difference, if any, a further 4 months had on how these women were feeling after their false-positive mammogram. This was also 1 month before those on early recall were due for another screening mammogram. In this prospective cohort study, the previous studies' results (Ong *et al*.[104]) were used as the baseline measures for comparison with the same subgroups. This study only included 12 centres, as one centre did not put any women on early recall.

Although this was a fairly good-quality study,[103] it did not quite meet the same standards of reporting as the previous one.[104] Omissions include not considering potential sources of bias, giving an explanation of how missing data were handled and not providing demographic information.

The latest publication from OPCERG on the population at normal risk of breast cancer is by Brett and Austoker[59] ($n = 505$). This study follows the same cohort of women as Brett *et al*.[103] in 13 NHSBSP clinics in England and Scotland. They took measures of adverse psychological consequences with the PCQ at 35 months after participants' last assessment (i.e. 1 month before their next routine screening mammogram was due). Brett and Austoker[59] deemed that a total score >12 (range 0–36) on the PCQ showed negative psychological consequences and so it is only the percentage of scores >12 that are reported. It is not clear why they have chosen this total score as the cut-off point for showing psychological harm. The original validation paper by Cockburn *et al*.[67] does not have this cut-off but indicates that the bottom quartile of scores represent no dysfunction, the second quartile mild psychological disturbance, the third quartile moderate disturbance and the top quartile marked disturbance.

Brett and Austoker[59] also used an ad hoc questionnaire to measure satisfaction with the breast screening service and assess factors that may influence women's level of anxiety. They also asked women about their intention to attend their next routine mammography screening. This last outcome was compared with their actual attendance.

Again, this is a generally well-described study but with similar omissions as before: no explanation of how missing data were dealt with or consideration of the role bias may have played in the results. However, they do provide some demographic information about the participants (marital status, home ownership and educational level).

Prior to the work by the Oxford team, Sutton *et al*.[55] ($n = 1021$) conducted a retrospective cohort study that was primarily interested in the levels of anxiety experienced by women attending routine mammography screening who had normal outcomes. However, they also looked at anxiety levels in women who had false-positive results. Nine months after the pre-screening baseline, they asked these women and others with normal results to retrospectively reflect on how anxious they had felt at six stages of the screening process: (1) receiving the invitation; (2) waiting at the clinic for the mammogram; (3) at

TABLE 6 Summary of key quality markers of included observational studies

STROBE statement – checklist of items that should be included in reports of observational studies																
	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Title and abstract																	
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	✗	✗	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Introduction																	
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Objectives	3	State specific objectives, including any pre-specified hypotheses	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Methods																	
Study design	4	Present key elements of study design early in the paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up and data collection	✓	✓	✓	✓	P	✓	✓	✓	✓	P	✓	✓	✓	✓	✓

STROBE statement – checklist of items that should be included in reports of observational studies	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Participants	6	(a) *Cohort study* – give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA
		Case–control study – give the eligibility criteria and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		Cross-sectional study – give the eligibility criteria and the sources and methods of selection of participants	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	✓	✓	✓	✓
		(b) *Cohort study* – for matched studies, give matching criteria and number of exposed and unexposed	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	P
		Case–control study – for matched studies, give matching criteria and the number of controls per case	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	✓

continued

TABLE 6 Summary of key quality markers of included observational studies (*continued*)

STROBE statement – checklist of items that should be included in reports of observational studies																
	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders and effect modifiers. Give diagnostic criteria, if applicable	✓	✓	✓	✓	P	P	✓	✓	✓	✓	✓	✓	✓	✓	P
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	✓	✓	✓	P	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bias	9	Describe any efforts to address potential sources of bias	✓	✗	✓	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗
Study size	10	Explain how the study size was arrived at	✓	✓	✓	✗	✓	✗	✗	✓	NA	NA	✗	✗	✗	✗	✗
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	✗

STROBE statement – checklist of items that should be included in reports of observational studies	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✗	✓	✓
		(b) Describe any methods used to examine subgroups and interactions	NA	NA	✓	NA	NA	NA	NA	NA	✗	NA	NA	NA	NA	NA	NA
		(c) Explain how missing data were addressed	✗	✗	✓	✗	✗	✗	✗	✗	NA	NA	✗	✗	NA	✗	NA
		(d) *Cohort study* – if applicable, explain how loss to follow-up was addressed	✗	✗	NA	✗	✗	✗	✗	✓	NA	NA	NA	NA	NA	NA	NA
		Case–control study – if applicable, explain how matching of cases and controls was addressed	NA	NA	NA	NA	NA	NA	✓	NA	NA	NA	NA	NA	NA	NA	NA
		Cross-sectional study – if applicable, describe analytical methods taking account of sampling strategy	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	✗	NA	NA	NA	✗
		(e) Describe any sensitivity analyses	NA	NA	NA	NA	NA	NA	✓	✓	NA	NA	NA	NA	NA	NA	NA

continued

TABLE 6 Summary of key quality markers of included observational studies (*continued*)

STROBE statement – checklist of items that should be included in reports of observational studies																
	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Results																	
Participants	13	(a) Report numbers of individuals at each stage of study, e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up and analysed	✓	✓	✓	✓	P	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		(b) Give reasons for non-participation at each stage	✓	✗	NA	✗	✗	✗	✗	✗	NA	NA	✗	✗	NA	NA	✗
		(c) Consider use of a flow diagram	✗	✗	NA	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗
Descriptive data	14	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	✓	✗	P	P	P	P	✓	✓	P	✗	✗	✗	NA	✗	P
		(b) Indicate number of participants with missing data for each variable of interest	✗	✗	✓	✗	✗	✗	✗	✓	NA	NA	✗	✗	NA	✓	✗
		(c) *Cohort study* – summarise follow-up time (e.g. average and total amount)	✓	✓	NA	✗	✗	✗	✗	✗	NA	NA	NA	NA	NA	NA	NA

STROBE statement – checklist of items that should be included in reports of observational studies

	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Outcome data	15	*Cohort study* – report numbers of outcome events or summary measures over time	✓	✓	NA	✗	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA
		Case–control study – report numbers in each exposure category, or summary measures of exposure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		Cross-sectional study – report numbers of outcome events or summary measures	NA	NA	✓	NA	NA	NA	NA	NA	NA	NA	✓	✓	✓	✓	✓
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% CI). Make clear which confounders were adjusted for and why they were included	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	P	✓	✓	✓	✓
		(b) Report category boundaries when continuous variables were categorised	NA	NA	NA	NA	✓	✓	NA	NA	NA	NA	NA	NA	NA	NA	NA
		(c) If relevant, consider translating estimates of RR into absolute risk for a meaningful time period	✗	✗	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

continued

TABLE 6 Summary of key quality markers of included observational studies (*continued*)

STROBE statement – checklist of items that should be included in reports of observational studies																
	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Other analyses	17	Report other analyses done (e.g. analyses of subgroups and interactions, and sensitivity analyses)	NA	✓	✓	✓	NA	NA	✓	✓	NA	NA	NA	NA	NA	NA	NA
Discussion																	
Key results	18	Summarise key results with reference to study objectives	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	✗	✓	✗	P	✗	✗	✓	✓	✗	✗	✗	✗	✗	P	✗
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies and other relevant evidence	✓	✓	P	P	✓	P	✓	✓	✓	✓	✓	✓	P	✓	P
Generalisability	21	Discuss the generalisability (external validity) of the study results	✗	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓	✓	✓	✓	✓

STROBE statement – checklist of items that should be included in reports of observational studies	Item no.	Recommendation	Brett and Austoker 2001[59]	Brett *et al.* 1998[103]	Ong *et al.* 1997[104]	Sutton *et al.* 1995[55]	Bull and Campbell 1991[106]	Ellman *et al.* 1989[105]	Brain *et al.* 2008[107]	Tyndel *et al.* 2007[101]	McCann *et al.* 2002[61]	O'Sullivan 2001[108]	Orton *et al.* 1991[109]	Ong and Austoker 1997[110]	Ong *et al.* 1996[111]	Austoker and Ong 1994[112]	Smith *et al.* 1991[113]
Other information																	
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗

✓, item present; ✗, item absent; F, results only present as figures; NA, not applicable; P, item partially present.

the clinic after the mammogram; (4) waiting for the results; (5) after reading the results letter; and (6) now. These measures were taken on a three-point ad hoc scale, ranging from not anxious (1), a bit anxious (2) to very anxious (3). The retrospective results are only reported numerically at stages 3–5, therefore these data have been extracted. The data at other time points are reported in graphical form only and so could not be reliably extracted.

The baseline measures were taken using different measures (STAI[76] and the GHQ[72]); however, the results of the women with false-positive mammograms are not disaggregated from the rest of the participants and so are not reportable.

This is a poorly designed and executed study. Initially, a cohort were recruited and followed up through the screening process. Additionally, another sample group was recruited which overlapped with the original sample group. A subset that contained participants from both groups was used for the main analysis (which compared women with normal mammograms with non-attenders and is therefore outside the scope of this review), while the whole of the first sample was used in the retrospective analysis of anxiety. There was no consideration of bias, confounding or the weakness of the design that relied on memory going back 9 months using an unvalidated measure. The results of the false-positive substudy are only partially numerically reported, the whole being shown in a figure that does not allow accurate disaggregation. Furthermore, there is no accounting of attrition or information about how missing data were dealt with. There is, however, acknowledgement that if they had used a disease-specific measure, such as the PCQ, they might have found different results.

Another similarly poor-quality study was conducted by Bull and Campbell[106] in 1991 ($n = 750$). They followed a cohort of women who had attended mammography screening from their invitation to screening until 6 weeks after they had received the results of their follow-up assessment. The women were divided into four groups: (A) those invited for screening; (B) those with a normal mammogram; (C) those who went for further assessment with a mammogram, ultrasound or FNA; and (D) those who had a surgical biopsy for further assessment. The sample in groups A and B overlapped. However, no rationale is given for the design of the study. The outcomes measured were anxiety and depression using HADS,[74] frequency of breast self-examination and impressions of the screening programme using an open question in an ad hoc questionnaire.

This study had a number of reporting flaws; in particular, there was no comment on the possibility of bias from the design and conduct of the study, nor was there any explanation of how missing data were accounted for. The only description of the participants was their age and no information was given about attrition or reasons for this. The discussion speculates at length about possible reasons for the findings, but neglects to mention any limitations of the study. Bull and Campbell[106] also fail to comment on the generalisability of their findings.

The prospective cohort study by Ellman *et al*.[105] ($n = 752$) is of similarly poor quality. Ellman *et al*.'s[105] aim was to compare the psychiatric morbidity experienced by women who had attended breast screening or a review clinic and had one of five results (group A, normal; group B, false-positive; group C, symptomatic with benign result; group D, symptomatic or recalled women who did have cancer; or group E, women with a history of breast cancer with or without symptoms). Only women in groups A and B were within the inclusion criteria of this systematic review and had their results reported. Morbidity was measured with GHQ-28[72] either at the breast screening clinic (groups A and B) or the review clinic (groups C–E) before participants saw the clinician and was measured again 3 months later. An ad hoc questionnaire seeking opinions about the clinic and their experience of treatment was given to participants at follow-up.

Ellman *et al*.'s study[105] failed to show an understanding of the limitations of their design. There was no consideration of potential sources of bias, no explanation of how missing data were accounted for or how this might affect the results. Although the tables of results were clearly presented and the results summarised in the discussion, there was no consideration of how generalisable they might be.

Psychological impact with family history of breast cancer

Although the population in this section is not part of the general breast cancer screening population or programme, it is interesting to look at this research and compare the results with that of the general population. The OPCERG and others collaborated in the Psychological Impact of Mammography Screening (PIMMS) research programme to look at the effects of having a false-positive screening mammogram in a population of women with a FHBC. The three papers reported in this section are from the PIMMS group.

Tyndel *et al*.[101] published the first paper in the series in 2007 ($n = 2321$). This was a prospective cohort study of women aged 35–49 years who had a moderate-to-high risk of breast cancer due to their family history and were on an annual screening programme. They were interested in comparing the psychological impact of having a false-positive mammogram with that of having a normal one within this group. Women who had previously had cancer and those with a family history of ovarian cancer were excluded. Disease-specific outcomes (PCQ and CWS-R[68]) were measured 1 month before participants' screening mammogram, 1 month after the 'all-clear' and 6 months after assessment.

This is a reasonably good-quality observational study. It is well-designed and well-reported; almost all the quality indicators in the STROBE checklist are met. The only substantial omission by Tyndel *et al*.[101] is a failure to report the extent of missing data and how these items were accounted for in the analysis.

The follow-on study is by Brain *et al*.[102] ($n = 1286$) who conducted a multiple regression analysis of the participants of the Tyndel *et al*. study[101] to find out which pre-screening variables were predictive of cancer-specific distress. Initially, they conducted partial correlations of pre-screening variables on a number of scales (CWS-R,[68] brief COPE,[71] perceived risk of cancer and dispositional optimism) to see how they correlated with cancer worry at 1 and 6 months after assessment. This was followed by a hierarchical multiple regression to show predictive associations between these baseline variables and cancer worry at 1 and 6 months.

This second study from the PIMMS group was also of reasonable quality. The main criticism is the same: a failure to report how many data were missing and how missing data were dealt with in the analysis.

A qualitative interview study was nested within the PIMMS research. However, this was published only as the summary of a conference poster. The summary is reasonably detailed and, although it does not give the level of detail that would be expected in a qualitative research paper, it has been included because it is part of an included study and it gives some insight into the results found by the quantitative research. This interview study was conducted by Clements *et al*.[107] on 58 women who were part of the PIMMS cohort study (false-positive result = 22, normal result = 36).

As this study was published only as the summary of a poster, many of the expected quality criteria are unreported. However, the design is suitable for the research question and the thematic method of analysis appropriate. There is a lack of detail about the results and their interpretation as well as a lack of a theoretical framework and justification for the methods used. Therefore, it is not possible to comment thoroughly on the robustness of this study.

Impact of false-positive mammograms on returning for routine screening

The studies in this section have been subdivided into those that report on participants' actual returning for future breast screening and those that report on the intention to reattend.

Actual reattendance

Five studies measured the reattendance at routine mammography screening of women who had a false-positive screening mammogram.

The most recent study to report the actual reattendance of women who had a false-positive screening mammogram was by McCann *et al*.[61] ($n = 140,387$). This was a large, retrospective study that looked at

records going back 3.5 years from NHS screening units in East Anglia and Cancer Registry databases. As well as looking at reattendance, this study aimed to quantify the increased risk of interval cancer (cancer detected between screenings) found among women in this group and establish if there was also an increased risk of cancer detection at second-round screening if women had been false-positive in the first round. The risk of interval cancer is outside the scope of this systematic review, therefore these results are not reported.

This large, convincing, retrospective cohort study gives clearly presented results that are well discussed. However, the statistical methods used are not described and the description of demographic characteristics is limited to age.

Prior to this, O'Sullivan *et al*.[108] ($n = 5649$) conducted a similar retrospective cohort study of women who had taken part in the NHS screening programme in east and central London. Their focus was only on the effects of having a false-positive mammogram on reattendance at the next screening round.

This data registry study clearly presents its descriptive results. However, O'Sullivan *et al*.[108] fail to provide any information about the characteristics of the participants, the number of screening centres involved or the possible role of bias and no comment is made about the study's generalisability to other screening centres.

The 2001 study by Brett and Austoker[59] ($n = 505$) looked at actual reattendance as well as the psychological impact of false-positive mammograms. A summary of this paper can be found above in *Primary research*.

Previously, Meldrum *et al*.[115] ($n = 3083$) conducted a RCT on reattendance that compared a standard invitation letter for second-round screening (3 years later) with a letter tailored to the outcome of the previous screening round (e.g. false-positive or normal). Additionally, a telephone interview study was conducted of 66 women who had tailored intervention letters to gauge their views of the letters.

This is a poor-quality RCT. It is not clear whether or not the participants were aware that they were taking part in a study. Very little detail is given about the intervention or the control letters, certainly not enough to replicate the study. Although the method of randomisation is reported, it is unclear whether or not the assessors were blinded to the groups. Most significantly, the analysis uses some inappropriate methods. In particular, an adjustment is made for multiple comparisons when only a single comparison is made. There is no explanation for this. This leads to an inaccurate interpretation of the results in the discussion, which fails to acknowledge any limitations of the study. The interview study is inadequately reported with no description of methods of analysis and what probably is very selective reporting of results. A CONSORT assessment of quality can be found in *Appendix 3*.

The final study in this group is by Orton *et al*.[109] ($n = 1582$) using data from three GP practices. They looked at whether or not the acceptability of the first round of screening or having a false-positive mammogram could help predict attendance at second-round screening 3 years later. Unfortunately, the authors did not disaggregate their data to show the effect of screening satisfaction for women with false-positive mammograms on attendance. The only usable data is a comparison of second-round attendance of women who had false-positive mammograms in the first screening round with those who had a normal result.

This retrospective cohort study reports most of the quality criteria for observational studies. The rationale and methods are clearly described. However, there is no description of the demographic characteristics of the participants, so it is not possible to judge whether or not the differences in outcomes may have been because of bias or confounding factors. The discussion fails to comment on the limitations of the study, although it does remark on the generalisability of the study.

Intended reattendance

The most recent research looking at the intention to reattend the next screening round following a false-positive mammogram is the 2001 study by Brett and Austoker.[59] This was not the main focus of the paper: the data consist of women's responses to a questionnaire about which external influences have affected their attitude to returning for screening in 1 months' time. It is not possible to compare this intentional data with the actual reattendance. A summary of this paper can be found above in *Psychological impact*.

In addition to measuring adverse psychological consequences at 5 months post assessment, Brett *et al*.[103] ($n = 284$) also measured participants' intention to attend their next breast screening appointment. For most of the cohort, this was 2.5 years away, but for those placed on early recall it was only 1 month away. A summary of this paper can be found above in *Psychological impact*.

Interventions to reduce the impact of being recalled for further assessment

No studies were found that directly addressed the issue of testing an intervention to relieve the negative psychological consequences of false-positive mammograms. However, two studies were found that looked at the information needs of women prior to further assessment, one from OPCERG (1994–7)[110–112] and one from Smith *et al*.[113] These studies do not disaggregate the data according to the outcome of the recall (i.e. distinguish between those who had false-positive and true-positive outcomes). This breaches the inclusion criteria for this systematic review. However, these studies are included because at the time of recall the final outcome is unknown to the women and there is no reason to suspect that, prior to the event, women with false-positive outcomes would have different information needs from those with other outcomes.

From 1994 to 1998 Austoker, Ong and others from OPCERG published a series of three papers about their research into the information needs of women who were recalled for further assessment following routine mammography screening, plus national guidelines on behalf of NHSBSP. This research will be presented first followed by a study by Smith *et al*.[113] of recall invitation letters.

The original work that OPCERG studies came from is by Austoker and Ong[112] ($n = 1493$). Their aim was to assess the need for written information among women who had been recalled for assessment following routine screening mammography. They did this by conducting a multicentre study of eight breast screening centres in the UK.

First, 484 women were invited for a questionnaire interview after recall and before assessment. Two weeks after assessment these and other women (total $n = 1493$) were given an ad hoc questionnaire to assess their response to being recalled and how worrying or reassuring the messages contained in recall literature were. Second, Austoker and Ong[112] also collected samples of the literature sent to women from the different centres and found a wide variation in the amount of information given and the language used. They assessed the breadth of information given in the recall letters and leaflets and whether or not the women thought they had been given sufficient information on matters such as the reason for recall, the location of the centre, who could come with them, how long the appointment would take, what tests they would undergo, who they would see, when they would know the outcome and how to get more information. They also assessed the language used in the information literature and categorised it according to whether the participants found it to be particularly worrying or reassuring. Simple descriptive statistics are reported together with a few examples of comments made in the open questions.

This is a well-presented and clearly reported study. However, Austoker and Ong[112] failed to gather information about the demographic profile of their participants; given that this was a large study, it would have been very interesting to see what light this information might have shone on the results.

Following on from this, and based on their original development of criteria for evaluating breast screening information material, Ong *et al*.[111] ($n = 84$) evaluated the health education literature in 84 breast screening units in the UK. They considered both the initial screening letter given to women and the recall letter and leaflet sent to those with suspicious mammograms. The evaluation criteria were divided into words and

phrases that were particularly worrying and those that were particularly stress relieving. The number of centres that mentioned items in their literature in these criteria was recorded.

As this is an evaluation of written material with simple descriptive statistics, many of the STROBE quality criteria do not apply. The methods and results were clearly described and comprehensively discussed. The main criticism is that there is a lack of discussion of the limitations of the study.

This study led into the most recent of the primary research papers, which is by Ong and Austoker[110] ($n = 1493$). They used an ad hoc questionnaire to elicit views of women about their experience of being recalled, in particular the quality of communication at the clinic (who was available to talk to at the clinic and related information needs). The questionnaire also had open questions that allowed a free response that underwent a discourse analysis.

Overall, this is a well-conducted and clearly reported survey. However, there are some omissions. The authors have failed to report the demographic characteristics of the participants or discussed what effect these may have had on the results, given no indication of how bias or confounding factors may have played a role in the results and have not considered other limitations of their study. Furthermore, the methods and results of the discourse analysis are not described, except to say that the most frequently occurring items or those deemed to be most important by the analysts are tabulated.

The most recent publication in this series is by Ong *et al.*[114] and is the national guidelines for improving the quality of written information sent to women who are recalled for assessment. The guidelines are based on the findings of the previous three papers and give advice on the content and wording of assessment invitation letters. The research evidence that is included in the report is a conflation of the findings of the studies and does not contain new evidence. Therefore, these data have not been extracted as they are already in the primary research paper data extraction forms. The guidelines are mentioned here for completeness.

The final study in this section is a poorly described survey of women's satisfaction with the Leicestershire Breast Screening Service in 1989–90 by Smith *et al.*[113] ($n = 103$). This includes assessing the adequacy of information given to women recalled for further assessment. Their aim was to compare three different versions of a recall letter giving different amounts of information, including telephone access to a breast care nurse, to see which was most acceptable to recalled women.

This study is described as an audit, but it goes further than that as it includes experimental testing of the different forms of the recall letter. The reporting of the study is very inadequate. Not enough information about the methods used is provided; it is often difficult to reconcile the text with the tables and there is no acknowledgement of the limitations of their methods.

Results and comment

Systematic reviews and meta-analyses

The scope of the included systematic reviews was broader than that of this review; they included non-UK studies and outcomes measured for < 1 month from women receiving the 'all-clear' after their follow-up assessment. Therefore their results may not reflect the medium- to long-term outcomes in the UK that are the subject of this review. Nevertheless, they are of interest as a comparison to this review's findings.

All the systematic reviews and meta-analyses summary results showed a negative impact from receiving a false-positive mammogram on measures of well-being, depression and anxiety compared with women with normal screening results. The primary studies included used both disease-specific and generic measures, which were usually analysed together in the reviews. The exception to this was the meta-analysis of generic psychological measures by Salz *et al.*[85] that showed no difference in psychological

outcomes between women with false-positive results and those with normal ones except for anxiety, which was positively correlated with having a false-positive mammogram [0.03 (95% CI 0.00 to 0.07)]. The evidence varied concerning whether psychological distress had a short-term (< 1 month after assessment) or long-term impact. There was some evidence that the degree of impact varied with the severity of the reassessment test, with women undergoing biopsy showing greater psychological distress than those with a repeat mammogram.[82]

The results for the impact of receiving a false-positive mammogram on self-care behaviour and returning for the next routine screening mammogram give a more complex picture. Two reviews, Salz *et al*.[85] and Armstrong *et al*.,[87] looked at effects of having a false-positive mammogram on frequency of breast self-examination. In both evidence syntheses, women with false-positive mammograms reported increased breast self-examination, which may indicate increased anxiety about the risk of developing breast cancer. However, Armstrong *et al*.[87] found there was no statistically significant difference between groups in the likelihood of returning for routine breast screening, although it is not clear whether the studies were reporting actual attendance or intention to attend. This may be important as Bankhead *et al*.[84] found that women were more likely to say that they had an intention to attend their next routine mammogram than actually do so. Other studies showed a variation in the effect of a false-positive mammogram on returning for screening according to location, with European women unaffected in this domain, Canadian women less likely to return and women from the USA more likely to return for routine mammography.[62] *Table 7* gives a summary of the results of the included systematic reviews and meta-analyses.

Primary research

As in the previous section, the results of the primary research studies have been organised according to their domain of interest. *Figure 5* (see *Primary research*) shows the relationship of the included primary studies to the outcome categories.

With the exception of the intervention study by Meldrum *et al*.,[115] all the included studies had observational designs. These designs are necessarily used when looking at the effects of a past event. However, results from observational studies should be interpreted with caution as bias and confounding may have influenced them. There may be systematic differences between the groups that do and the groups that do not have false-positive mammograms: those in the false-positive group will have had greater exposure to NHS care which could have affected their response to the questionnaires, none of the studies were blinded, therefore the way that measures were taken may have varied between the groups, and attrition bias may have skewed the results either way, as those who withdrew from the study may have been more or less affected by their screening outcome than those who remained in the studies.

As reported above, some of the studies included participants who are outside the inclusion criteria for this systematic review. Therefore, only data from the study population included in this review have been extracted and reported.

Psychological impact

In the first paper from OPCERG, by Ong *et al*.[104] ($n = 877$), adverse psychological consequences were measured at T1 (1 month after the last screening appointment) using the negative PCQ subscale. The proportion of participants within each of the screening outcomes was recorded. The results are presented simply as proportions; therefore, the review authors (MB) have calculated RRs to provide greater insight into the relationship between false-positive and normal mammograms with adverse psychological consequences. *Table 8* shows the outcomes at T1. These results show an increased risk of psychological distress at 1 month after the last screening appointment for women who had a false-positive result compared with women who had a normal mammogram. The risk of distress increases in line with the intrusiveness of the assessment process, so that women who had another mammogram had a RR of 1.71 (95% CI 1.24 to 2.35), whereas women who had a biopsy were at the greatest risk, RR 2.96 (95% CI 2.19 to 4.01). Those put on early recall or who had a FNA also showed increased distress, RR 2.13 (95% CI 1.58 to 2.87) and RR 1.97 (95% CI 1.44 to 2.69), respectively.

TABLE 7 Summary of false-positive results of included systematic reviews and meta-analyses

Author	No. of included false-positive studies	Total false-positive study population	Results summary: effects of false-positive mammograms	Direction of change: Well-being impact	Direction of change: Behavioural impact
Salz *et al.* 2010[85]	17	20,781	Disease-specific measures detected associations with increased psychological distress and increased breast care-related activity on all eight outcomes. However, generic measures only detected an effect on one outcome (generalised anxiety)	↓[a] _[b]	↓
Hafslund and Nortvedt 2009[86]	13	14,530	Having a false-positive result had a negative impact on women's well-being and quality of life that continued for weeks or months after assessment	↓	Not measured

Outcome	Breast cancer-specific: Correlation coefficient, *r* (95% CI)	Generic: Correlation coefficient, *r* (95% CI)
Distress	0.16 (0.10 to 0.22)	–0.04 (–0.08 to 0.01)
Somatisation	0.12 (0.05 to 0.19)	0.01 (–0.09 to 0.11)
Depression		0.00 (–0.04 to 0.04)
Fear	0.08 (0.03 to 0.14)	0.11 (–0.18 to 0.39)
Anxiety	0.22 (0.18 to 0.27)	0.03 (0.00 to 0.07)
Worry	0.12 (0.08 to 0.16)	0.02 (–0.04 to 0.08)
Perceived likelihood of cancer	0.09 (0.04 to 0.14)	
Perceived benefits of mammography	0.11 (0.06 to 0.17)	
Frequency of breast self-examination	0.11 (0.04 to 0.19)	

Author	No. of included false-positive studies	Total false-positive study population	Results summary: effects of false-positive mammograms	Direction of change	
				Well-being impact	Behavioural impact
Brewer *et al.* 2007[62]	23	313,967	Women who had a false-positive mammogram were more distressed and anxious than those who did not. A meta-analysis of the effect on likelihood of returning for routine screening varied with location: women from Europe showed a non-significant trend towards not returning, risk ratio 0.97 (95% CI 0.93 to 1.01), Canadian women were less likely to return, risk ratio 0.63 (95% CI 0.50 to 0.80) and women from the USA were more likely to return, risk ratio 1.07 (95% CI 1.02 to 1.12), following a false-positive mammogram	↓	~[c] ↓[d] ↑[e]
Armstrong *et al.* 2007[87]	22	167,969	Overall, there was a short-term association between anxiety and depression and having a false-positive mammogram. However, some studies found this to be prolonged and recurring prior to the next routine screening. There was no increased or decreased likelihood of returning for screening among women who had a false-positive mammogram compared with those with a normal one. Although there was an increase in breast self-examination and breast- and non-breast-related health-care visits	↓	–[f] ↓[g]
Brett *et al.* 2005[82]	29	17,394	In general, there was a negative psychological impact on women who had a false-positive mammogram compared with those with normal mammograms. The greatest effect was found with those who had undergone biopsy. There was conflicting evidence about the duration of the impact and what the effects were on the likelihood of women returning for routine mammography	↓	↓ ↑
Bankhead *et al.* 2003[84]	13	40,495	There was a difference between women's stated intention to reattend for routine mammography following a false-positive mammogram and their actual attendance, which was lower than had been said	Not reported	↓

↓, results indicate a significant negative effect on the domain; ↑, results indicate a significant positive effect on the domain; ~, results show a non-statistically significant trend; –, no difference between groups.

a Breast cancer-specific measures.

b Generic measures.

c Europe.

d Canada.

e USA.

f No effect on returning for routine mammography.

g Increased self-examination and health-care visits.

The next paper in this series is by Brett *et al*.[103] ($n = 284$) and reports outcomes 5 months after the last screening appointment (T2). Although all the women included at T2 were respondents at T1, it appears that they are a selected subset of Ong *et al*.'s[104] participants, who were matched with T1 respondents placed on 6 months early recall, using undisclosed criteria. A comparison between T1 results in Brett *et al*.[103] (the selected subgroup) and T1 in Ong *et al*.[104] indicates that the Brett *et al*.'s[103] subset has a higher proportion of adverse PCQ scores than the original study found (compare *Table 8* with column 2 of *Table 9*), thus leading to higher RR estimates in the study by Brett *et al*.[103] than in that by Ong *et al*.[104]

There is another source of uncertainty about the comparability of Brett *et al*.'s T1 results[103] with Ong *et al*.'s T1 results:[104] it is unclear what the proportion of women experiencing adverse psychological events represents. In the earlier paper (Ong *et al*.[104]) the proportions represent those experiencing at least one psychological event on the PCQ and in the last paper in the series (Brett and Austoker[59]) it is the proportion of women scoring >12 on the PCQ. However, Brett *et al*.[103] do not disclose the method used for calculating the proportion. One clue is that they report that the PCQ is scored 1–4 rather than the conventional 0–3. If they have transposed the scores in this way, then a score >12 would be equivalent

TABLE 8 Adverse psychological consequences 1 month (T1) after the last screening appointment

Outcome of screening	Women reporting adverse PCs at T1 *n/N* (%)	RR[a] (95% CI)
Normal mammogram	38/130 (29)	Baseline
False-positive after		
Further mammography	64/128 (50)	1.71 (1.24 to 2.35)****
FNA	61/106 (58)	1.97 (1.44 to 2.69)****
Biopsy	26/30 (87)	2.96 (2.19 to 4.01)****
ER after assessment	81/130 (62)	2.13 (1.58 to 2.87)****

****$p < 0.0001$; ER, early recall at 6 or 12 months; PC, psychological consequence.
a RR calculated by authors.
Source: adapted from Ong *et al*.[104]

TABLE 9 Brett *et al*.:[103] comparison of adverse psychological consequences 1 month (T1) and 5 months (T2) after the last screening appointment

Outcome of screening	Women reporting adverse PCs at T1 *n/N* (%)	RR[a] (95% CI)	Women reporting adverse PCs at T2 *n/N* (%)	RR[a] (95% CI)
Normal mammogram	9/52 (17)	Baseline	5/52 (10)	Baseline
False-positive after				
Further mammography	29/51 (57)	3.29 (1.73 to 6.23)***	23/51 (45)	4.69 (1.93 to 11.38)***
FNA	26/41 (63)	3.66 (1.93 to 6.93)****	18/41 (44)	4.57 (1.85 to 11.26)***
Biopsy	21/23 (91)	5.28 (2.87 to 9.68)****	14/23 (61)	6.33 (2.59 to 15.50)****
ER after assessment	32/46 (70)	4.02 (2.15 to 7.50)****	27/46 (59)	6.10 (2.56 to 14.54)****

*$p < 0.001$; ****$p < 0.0001$; ER, early recall at 6 months; PC, psychological consequence.
a RR calculated by authors.
Source: adapted from Brett *et al*.[103]

to a score of at least 1 on the original paper (Ong *et al*.[104]) where the PCQ is scored 0–3. Assuming that the scores are directly comparable in this way it can be seen that, for all screening outcome categories, the level of psychological adverse effects has reduced over time (T1–T2) for the matched subset (see *Table 9*). This suggests that the matched subset in Brett *et al*.[103] is not representative of the population in Ong *et al*.[104] which would introduce a potential bias leading to overestimates of adverse psychological consequences at T1 for women having false-positive mammograms.

Furthermore, there is disagreement between tables within the paper by Brett *et al*.[103] about the benign biopsy and early recall results. In their table 1, 61% of women placed on early recall and 59% of women who had a biopsy show adverse psychological consequences, but in their table 2 these figures are reversed, although we are told that the difference between the scores at T1 and T2 was lower for women placed on early recall than for those who had a biopsy. Therefore, it has been assumed that table 2 is correct. The authors were contacted about this but no reply was received.

At T2 the results show that for all groups of women the level of distress had reduced. The greatest decrease in distress was for those women who had a biopsy at assessment (–30%). However, the RR of distress compared with those with normal mammograms had increased for all false-positive subgroups. Those women who had a biopsy continued to be at a greater risk of distress than those with a normal mammogram (RR 6.33; 95% CI 2.59 to 15.50). Those women who were put on 6 months early recall were also at a high RR (6.10; 95% CI 2.56 to 14.54). This is probably because they were only 1 month away from their next screening appointment (see *Table 9*).

Brett *et al*.[103] also conducted a logistic regression to find out which personal characteristics might be influencing the results. This showed that only having psychological consequences at 1 month (OR 5.82; 95% CI 2.70 to 12.56) and the kind of further investigation that women had (OR 4.40; 95% CI 1.35 to 14.35) were related to having psychological consequences at 5 months (*Table 10*).

In the final paper in the series, Brett and Austoker[59] ($n = 505$) report outcomes at 35 months (T3) from the last screening appointment and compare them with the same participants at T1. It appears that this cohort has not been matched, as in the previous paper by Brett *et al*.,[103] but is the same cohort as in the first paper by Ong *et al*.[104] However, in this last paper, cases are those with a PCQ score >12. If the assumption of comparability is made between methods used to calculate T1 and T3 scores, then the proportion of women with normal mammograms experiencing psychological adverse events is similar at T1 (26%) and T3 (25%), although these results are higher than at T2 (10%). For those women with false-positive

TABLE 10 Logistic regression: variables related to psychological consequences at 5 months (T2) after the last breast screening appointment

Variable	OR (95% CI)
Psychological consequences at 1 month	5.82 (2.70 to 12.56)***
Result group (type of investigation)	4.40 (1.35 to 14.35)**
Age of women	1.00 (0.98 to 1.03) NS
Apprehensiveness about attending	0.92 (0.80 to 1.07) NS
Greater perceived likelihood of ever getting breast cancer compared with the average woman	0.91 (0.35 to 2.34) NS
Likelihood of attending future breast screening	0.61 (0.03 to 11.93) NS
Need to discuss breast screening with someone	0.50 (0.24 to 1.02) NS

**$p<0.01$; ***$p<0.001$; NS, not significant.
Source: adapted from Brett *et al*.[103]

mammograms (other screening outcomes) the proportion experiencing adverse outcomes has reduced over time, although for those who were assessed with FNA the proportion who were distressed at T3 (45%) is similar that at T2 (44%) (*Table 11*).

An examination of the RR of experiencing psychological consequences at T3 and T1 show that this risk has diminished substantially for those assessed with a further mammogram, RR 1.28 (95% CI 0.82 to 2.00), as they were not statistically significantly more distressed than women who had a normal mammogram. Also, women who were placed on 6-months early recall showed a 23% reduction in their RR of distress at T3, RR 1.82 (95% CI 1.22 to 2.72). However, women who were assessed by FNA had the least reduction in RR (0.28) with a 35-month RR of 1.80 (95% CI 1.17 to 2.77). Those who had a biopsy maintained their status of having the highest RR of distress at T3, RR 2.07 (95% CI 1.22 to 3.52). *Table 11* shows the results comparing T1 and T3.

The results from the OPCERG study indicate that there is an enduring relationship, lasting at least 3 years, between having a false-positive mammogram and exhibiting negative psychological consequences. This effect is shown to be in proportion to the degree of invasiveness of the assessment procedure, with 52% of women who were assessed by biopsy, 45% assessed by FNA, 32% assessed by further mammogram and 46% of those placed on 6 months early recall, compared with 25% of those with a normal mammogram, still experiencing distress 3 years after their last screening appointment. The numbers of distressed women may partly be explained by the impending date for their next screening appointment, although the proportions are still likely to reflect their assessment procedure.

The forest plot in *Figure 6* gives an overview of the RR of psychological distress from a false-positive mammogram in relation to time and the method of assessment when compared with a normal mammogram.

The results from the ad hoc questionnaire, given at 35 months (T3), reported by Brett and Austoker[59] shed some light on factors that may have influenced the level of distress experienced. *Table 12* shows the items that were statistically significantly correlated with psychological distress at 1 month before the next routine screening (T3). The highest correlation with distress is with the lack of opportunity to talk to someone after the screening appointment ($r = 0.35$). The other most highly correlated factors are the waiting time between screening and assessment ($r = 0.30$), communication problems at the screening appointment

TABLE 11 Brett and Austoker:[59] comparison of adverse psychological consequences 1 month (T1) after last breast screening appointment and 1 month before the next screening (T3, 35 months after the first appointment)

Last breast screening results group (1995)	Women reporting adverse PCs at T1 *n/N* (%)	RR[a] (95% CI)	Women reporting adverse PCs at T3 *n/N* (%)	RR[a] (95% CI)
Normal mammogram	26/99 (26)	Baseline	25/99 (25)	Baseline
False-positive after				
Further mammography	47/93 (51)	1.92 (1.31 to 2.83)***	30/93 (32)	1.28 (0.82 to 2.00) NS
FNA	36/66 (55)	2.08 (1.40 to 3.09)***	30/66 (45)	1.80 (1.17 to 2.77)***
Biopsy	15/21 (71)	2.72 (1.78 to 4.17)****	11/21 (52)	2.07 (1.22 to 3.52)**
ER after assessment	62/100 (62)	2.36 (1.64 to 3.40)****	46/100 (46)	1.82 (1.22 to 2.72)**

***$p < 0.001$; ****$p < 0.0001$; ER, early recall at 6 months; NS, not significant; PC, psychological consequence.

a RR calculated by authors.

Source: adapted from Brett and Austoker.[59]

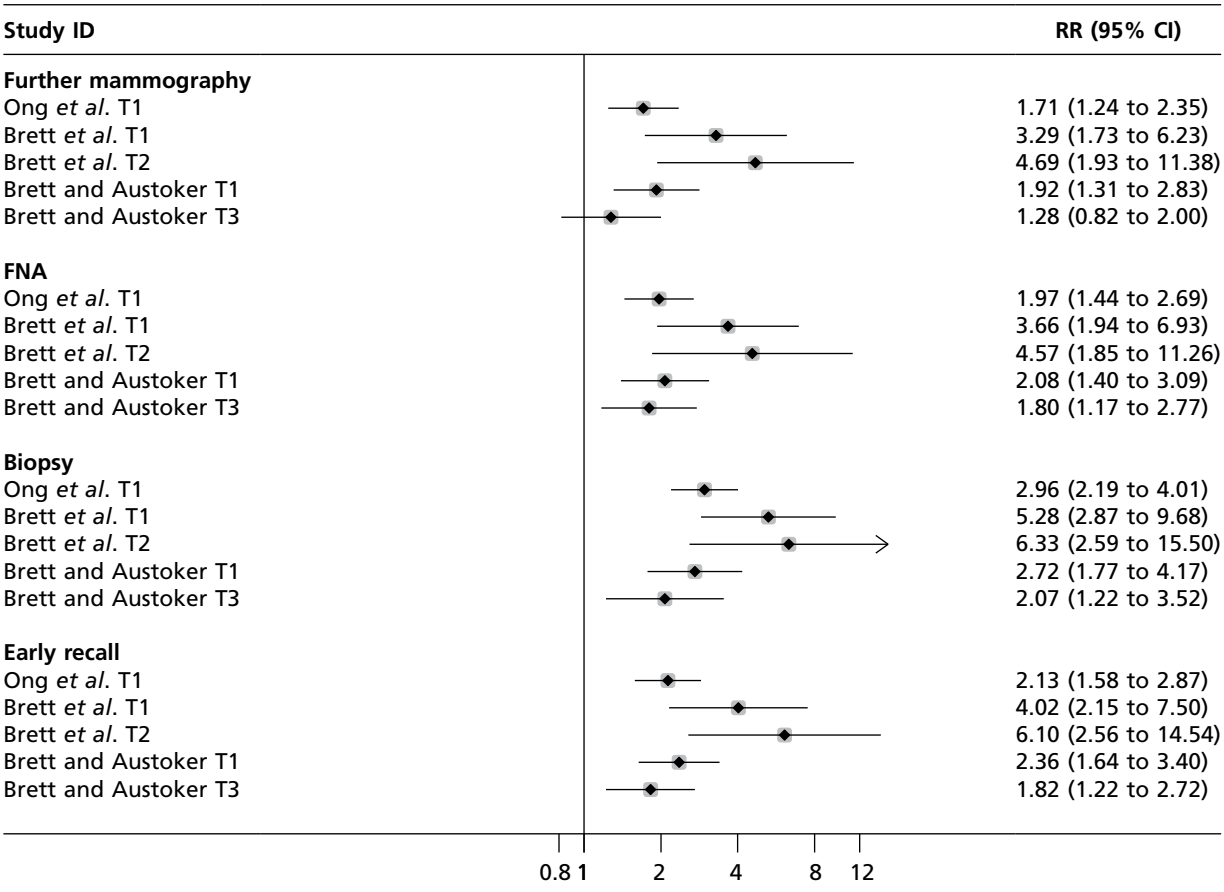

FIGURE 6 Forest plot of the RRs of negative psychological consequences from having a false-positive mammogram compared with a normal one by type of false-positive assessment at T1 (1 month after assessment), T2 (5 months after assessment) and T3 (35 months after assessment), measured with the PCQ.

due to anxiety ($r = 0.29$), fear of radiation ($r = 0.28$), amount of written information ($r = 0.28$), the performance of health workers ($r = 0.27$) and the unnecessary worry caused by the last breast screen ($r = 0.26$).

Preceding the work of OPCERG, Sutton *et al.*[55] ($n = 1021$) published results from their retrospective cohort study. These results show that, on their ad hoc questionnaire, when looking back over the previous 9 months, women who had false-positive mammograms were more likely to report that they were more anxious than women with normal mammograms at three time points: at the clinic after the mammogram, mean difference 0.24 (95% CI 0.03 to 0.45); while waiting for the results letter, 0.25 (95% CI 0.02 to 0.48); and after reading the results letter, 1.69 (95% CI 1.54 to 1.84). Unfortunately, Sutton *et al.*[55] do not numerically report how anxious the two groups were at the time they completed the questionnaire; these results are only presented as a graph which it is not possible to accurately read. The findings from this study are questionable as, apart from other methodological considerations of bias, the experience of having a false-positive mammogram may have coloured the women's view of how they felt at the time of screening and confounded their responses. It would seem unlikely that there should be any genuine difference between the two groups prior to receiving the results letter when it might be expected to find a difference in anxiety levels. A summary of the results from this study can be found in *Table 13*.

The results of Bull and Campbell's ($n = 750$) cohort study[106] are ambiguous. The authors' ad hoc questionnaire of the frequency of breast self-examination found that 6 weeks after receiving the 'all-clear' women who had a false-positive mammogram examined their breasts more frequently than women who had a normal mammogram. Women who had a biopsy at assessment were more likely to self-examine

TABLE 12 Correlation between psychological distress at 1 month before the next routine breast screening and dissatisfaction and anxiety related to past routine breast screening

Statements about last screening appointment	False-positive coefficient
Opportunity to talk to somebody after the breast screening appointment	0.35***
Waiting between appointment letter and appointment(s)	0.30***
Difficulties with taking in verbal information at breast screening appointment because of anxiety	0.29***
Fear of radiation	0.28***
Amount of written information	0.28***
Perceived performance of health workers	0.27***
Unnecessary worry experienced as a result of the last breast screening	0.26***
The amount of time spent on verbal communication at assessment	0.24***
Verbal communication: chance to say what is on one's mind	0.23***
Waiting for test results	0.22***
Quality of verbal communication	0.21***
Women's understanding of test result	0.21***
Amount of information provided in advance	0.18**
Postal notification of mammographic results	0.14*

*$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.
Source: adapted from Brett and Austoker.[59]

TABLE 13 Retrospective anxiety 9 months after receiving an invitation for breast screening

Outcome	Stage 1: receive screening invitation, mean (SD)	Stage 2: while waiting for the mammogram, mean (SD)	Stage 3: at the clinic after the mammogram, mean (SD)	Stage 4: after screening and before receiving the results, mean (SD)	Stage 5: afterreading the results letter, mean (SD)	Stage 6: now,mean (SD)
False-positive ($n = 24$)	Not reported	Not reported	1.60 (0.68)	1.95 (0.09)	2.85 (0.37)	Not reported
Normal mammogram ($n = 671$)	Not reported	Not reported	1.36 (0.52)	1.70 (0.57)	1.16 (0.36)	Not reported
Difference in means (95% CI)	–	–	0.24 (0.03 to 0.45)	0.25 (0.02 to 0.48)	1.69 (1.54 to 1.84)	–

Source: adapted from Sutton *et al.*[55]

once a week or more than those who had less invasive assessments. This may indicate that the women with false-positive results were more anxious about having breast cancer than those with normal results, although there is uncertainty about the meaning of very frequent breast self-examination. These results can be seen in *Table 14*.

When the results of both false-positive groups are combined there is a clear link between frequency of breast self-examination and having a false-positive mammogram, with the RR increasing with the frequency of self-examination (*Table 15*).

However, Bull and Campbell's[106] results with the HADS tell a different story. They show no statistically significant difference in anxiety or depression between the groups at 6 weeks after assessment, with the majority in all groups being within the normal range on both subscales. Borderline anxiety and depression were higher in the false-positive group that did not have a biopsy, depression (12%) or anxiety (12%) than in the false-positive group that had a biopsy, depression (6%) or anxiety (8%). This may have been because a biopsy was seen as a more conclusive declaration of health than other methods of assessment (*Table 16*).

When the false-positive groups are combined, the RR of experiencing anxiety or depression is not statistically significantly greater in the false-positive group (*Table 17*).

TABLE 14 Frequency of breast self-examination by group

Frequency of breast self-examination by group	Normal mammogram (*N* = 102), *n* (%)	False-positive (not biopsy) (*N* = 204), *n* (%)	False-positive (biopsy) (*N* = 49), *n* (%)
Never	22 (22)	24 (12)	7 (14)
Less than once a month	23 (23)	34 (17)	7 (14)
Once a month	47 (46)	97 (48)	18 (37)
Once a week	10 (10)	41 (20)	12 (24)
More than once a week	0	8 (4)	5 (10)
No response	0	0	0

Source: adapted from Bull and Campbell.[106]

TABLE 15 Relative risk of frequency of breast self-examination

Frequency of breast self-examination	Normal mammogram (*N* = 102), *n* (%)	All false-positive (*N* = 253), *n* (%)	RR[a] comparing normal mammogram with all false-positive (95% CI)
Never	22 (22)	31 (12)	0.57 (0.35 to 0.93)*
Less than once a month	23 (23)	41 (16)	0.71 (0.46 to 1.13) NS
Once a month	47 (46)	115 (45)	0.99 (0.77 to 1.27) NS
Once a week	10 (10)	63 (25)	2.54 (1.36 to 4.75)**
More than once a week	0	13 (5)	(0.66 to 182.48) NS
No response	0	0	

*$p < 0.05$; **$p < 0.001$; NS, not significant.
a RR calculated by authors.
Source: adapted from Bull and Campbell.[106]

TABLE 16 Hospital Anxiety and Depression Scale results by group

HADS subscales	Normal mammogram *n/N* (%)	False-positive (not biopsy) *n/N* (%)	False-positive (biopsy) *n/N* (%)
Depression			
Normal (0–7)	95/104 (91)	168/202 (83)	43/49 (88) NS
Borderline (8–10)	7/104 (7)	25/202 (12)	3/49 (6) NS
Abnormal (>10)	2/104 (2)	9/202 (5)	3/49 (6) NS
Anxiety			
Normal (0–7)	91/103 (88)	174/202 (86)	42/49 (86) NS
Borderline (8–10)	10/103 (10)	24/202 (12)	4/49 (8) NS
Abnormal (>10)	2/103 (2)	4/202 (2)	3/49 (6) NS

NS, not significant.
Source: adapted from Bull and Campbell.[106]

TABLE 17 Relative risk of anxiety and depression for women with false-positives measured by the HADS

HADS subscales	Normal mammogram *n/N* (%)	False-positive *n/N* (%)	RR[a] (95% CI)
Depression			
Normal (0–7)	95/104 (91)	211/251(84)	Baseline
Borderline (8–10)	7/104 (7)	28/251 (11)	1.71 (0.77 to 3.78) NS
Abnormal (>10)	2/104 (2)	12/251 (5)	2.61 (0.60 to 11.44) NS
Anxiety			
Normal (0–7)	91/103 (88)	216/251 (86)	Baseline
Borderline (8–10)	10/103 (10)	28/251 (11)	1.16 (0.59 to 2.30) NS
Abnormal (>10)	2/103 (1)	7/251 (3)	1.46 (0.31 to 6.90) NS

NS, not significant.
RR calculated by authors
Source: adapted from Bull and Campbell.[106]

The prospective cohort study by Ellman *et al*.[105] ($n = 752$) used the GHQ-28,[72] another generic instrument, to measure psychological morbidity at the screening visit and 3 months later. They found that there were no statistically significant differences between women who had normal mammograms and those with false-positive ones. Scores >4 on the GHQ-28 are deemed to be 'cases' (*Table 18*).

The distribution of scores on the four subscales of somatic, anxiety, social dysfunction and depression can be seen in *Table 19*. These show that the proportion of scores above 'case' level declined in both groups between baseline and follow-up, with the greatest decline in the false-positive group for the somatic and anxiety subscales (40% to 26% and 44% to 29%, respectively).

When the RRs are calculated for psychological morbidity these again show that there is no statistically significant difference between those women with and those without false-positive mammograms at baseline and 3 months later (*Table 20*).

TABLE 18 Proportion of GHQ scores (>4) showing psychological morbidity at the screening clinic and 3 months later

Time	Normal mammogram *n*/N (%)	False-positive *n*/N (%)
Screening visit	71/295 (24)	216/721 (30) NS
3 months later	56/295 (19)	137/721 (19) NS

NS. not significant.
Source: adapted from Ellman *et al*.[105]

TABLE 19 Distribution of GHQ-28 subscale scores showing percentage of cases

	Normal mammogram		False-positive	
Symptom subscale	Screening visit *n*/*N* (%)	3 months later *n*/*N* (%)	Screening visit *n*/*N* (%)	3 months later *n*/*N* (%)
Somatic	113/295 (38)	98/287 (34)	108/271 (40)	69/266 (26)
Anxiety	104/295 (35)	75/287 (26)	119/271 (44)	77/266 (29)
Social dysfunction	104/295 (35)	86/287 (30)	89/271 (33)	77/266 (29)
Depression	42/295 (14)	29/287 (10)	38/271 (14)	27/266 (10)

Source: adapted from Ellman *et al*.[105]

TABLE 20 Relative risk of psychological morbidity GHQ-28

Time	Normal mammogram *n*/*N* (%)	False-positive *n*/*N* (%)	RR[a] (95% CI)
Screening visit score			
Normal (0–4)	222/295 (75)	189/271 (70)	Baseline
Subclinical/mild (5–9)	49/295 (17)	48/271 (18)	1.12 (0.78 to 1.60) NS
Clinical mild/moderate (10–28)	24/295 (8)	34/271 (13)	1.56 (0.96 to 2.55) NS
3 months later score			
Normal (0–4)	232/287 (81)	216/266 (81)	Baseline
Subclinical/mild (5–9)	31/287 (11)	23/266 (9)	0.82 (0.49 to 1.36) NS
Clinical mild/moderate (10–28)	24/287 (8)	27/266 (10)	1.19 (0.70 to 2.00) NS

NS, not significant.
a RR calculated by authors.
Source: adapted from Ellman *et al*.[105]

The forest plot in *Figure 7* shows the results from Bull and Campbell[106] and Ellman *et al*.[105] Here it can be seen that, although none of the results show clinical levels of general anxiety and depression, there is a trend in that direction.

Summary: psychological impact in the general population

The OPCERG found that there was a statistically significant negative psychological impact from having a false-positive mammogram, whereas the main outcome measures of Bull and Campbell[106] and those of Ellman *et al*.[105] did not. The difference in findings between those of OPCERG and others may be explained in a variety of ways by differences in their design, methods and populations or they may be an artefact of

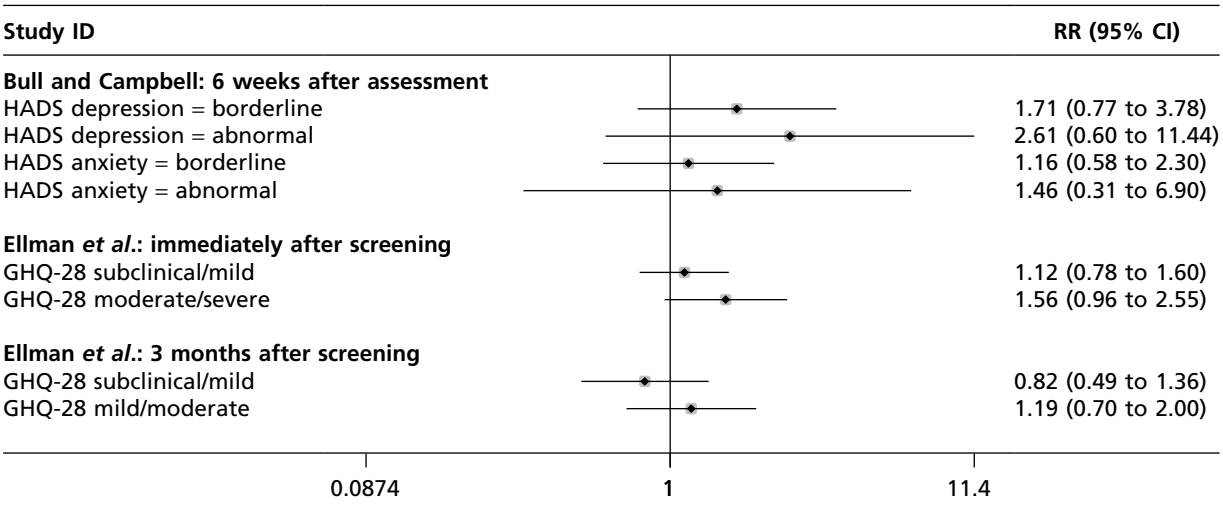

FIGURE 7 Relative risk of suffering clinical levels of general anxiety and depression following a false-positive mammogram compared with a normal mammogram, measured by Bull and Campbell[106] (HADS) and Ellman *et al*.[105] (GHQ-28).

bias or confounding. However, one key difference stands out: OPCERG measured outcomes with disease-specific instruments, whereas Bull and Campbell[106] and Ellman *et al*.[105] used generic measures designed to detect general anxiety and depression at clinically recognisable levels.

Further possible evidence of some distress comes from the Bull and Campbell study's[106] frequency of breast self-examination results. In this study there is a clear relationship between the RR of increasingly frequent breast self-examination and having a false-positive mammogram, with those women assessing themselves once a week or more than this having RRs of 2.54 (95% CI 1.36 to 4.75) and 10.95 (95% CI 0.66 to 182.48), respectively. It may be reasonable to suggest that this level of frequency of self-examination is a proxy for anxiety.

The disagreement about the psychological impact of false-positive screening results may be explained by the type of outcome measures used, whether disease-specific or generic. Brodersen *et al*.,[116–119] from the University of Copenhagen, have written a number of papers on the subject of measuring psychological distress in women who have received false-positive mammograms. In particular, they conducted a literature review to find out how suitable the outcome measures used in studies of false-positive mammograms were for detecting psychological distress.[116] The review found 23 includable studies; the most commonly used outcome measures were the HADS, GHQ, STAI and PCQ. By judging the instruments' psychometric properties in the context of false-positive mammography, they found that HADS, GHQ and STAI were unsuitable for use in measuring psychological distress. This was because the content of HADS and the GHQ were not applicable to the screening context as they were designed to screen for general anxiety at clinical levels rather than the specific breast cancer anxiety that may be expected. The STAI had been validated using students before their exams and may not replicate the anxiety felt by women possibly facing cancer. None of these scales have had their content validity demonstrated in the breast cancer screening context and all contain items that are irrelevant to women being screened, which may lead to the items' omission by respondents. Brodersen *et al*.[116] also reported that the PCQ was the most suitable measure to use to test psychological morbidity resulting from breast cancer screening as it had been validated in this context and was specifically designed for this purpose. However, its ability to reliably measure the psychological consequences of mammography screening in the long term has yet to be determined. *Table 21* shows the relationship between the types of measure used (generic or disease-specific) and the results of studies that compared women with false-positive mammograms with those with normal mammograms.

Therefore, it may be reasonable to speculate that, for those in the general population, having a false-positive screening mammogram can cause breast cancer-specific psychological distress that may endure for up to 3 years. However, it is less likely that there will be general anxiety detectable at clinically recognisable levels. Further research is needed in well-designed observational studies that use both disease-specific and generic outcomes to determine whether or not this is the case.

Psychological impact with family history of breast cancer

The PIMMS Management group's first results are from Tyndel *et al.*'s[101] ($n = 2321$) study in the FHBC population. When their within-study group results are considered, the women who had a normal mammogram showed a statistically significant decrease in their distress levels between T1 (1 month before screening) and T2 (1 month after screening) on both measures (CWS-R and PCQ), whereas those in the recall group did not. However, when the between T2-and-T3 (6 months after screening) scores are compared, both groups show statistically significant reductions in distress over this 5-month period on both measures (*Table 22*).

The between-group scores are harder to interpret owing to the potential for bias to be introduced by lack of randomisation and demographic data that indicates that there are differences between the groups. The false-positive group have statistically significantly greater proportions of participants with biological children ($p = <0.05$), a high-risk family history ($p = <0.05$) and post-mammography hospital attendance for symptoms ($p = <0.05$). The only statistically significant difference in the between-group scores is with the PCQ at T2, when those in the recall group {mean [standard deviation (SD)] 7.1 (7.44)} were more distressed than those in the normal result group [mean (SD) 4.08 (6.19)], this difference was no longer statistically significant 5 months later at T3 (see *Table 22*). However, these results, like those from the other studies, have not been adjusted for potential confounders (items on the PCQ and CWS-R). Tyndel *et al.*[101] present adjusted results (for the potential confounders listed above) in graphical form only (which could not be accurately transposed) that indicate that if these are considered then the recall group showed the greatest decrease in distress between T2 and T3 and the normal result group between T1 and T2 on both measures. The adjusted results were statistically significant; the level of significance is not reported.

Unusually, Tyndel *et al.*[101] also took measures on the positive subscale of the PCQ to see if there were any benefits from having a false-positive mammogram. They found that women who were recalled scored statistically significantly more highly at T2 [mean (SD) 13.02 (7.6)] than those with normal mammograms [mean (SD) 10.81 (6.9)]. However, this effect had diminished and was not statistically significant by T3. Additionally, when they had received their results at T2 and T3, women were asked whether their opinion of the benefits of breast screening had changed since their last visit. Women who were recalled were statistically significantly more likely to feel positive about the benefits of screening at both time points than those with normal mammograms, T2 = OR 3.16 (95% CI 2.14 to 4.70) and T3 = OR 2.35 (95% CI 1.53 to 3.61) (*Table 23*).

TABLE 21 Comparison of studies that used both disease-specific or generic outcomes and their results

		Disease-specific	Generic		Negative psychological consequences: false-positive vs normal result	
Study	Year	PCQ	HADS	GHQ-28	Found	Not found
Brett and Austoker[59]	2001				✓	
Brett *et al.*[103]	1998				✓	
Ong *et al.*[104]	1997				✓	
Bull and Campbell[106]	1991					✓
Ellman *et al.*[105]	1989					✓

TABLE 22 Comparison of distress at T1 (1 month before screening), T2 (1 month after screening) and T3 (6 months after screening)

Questionnaire	False-positive result	Within false-positive result	Normal result	Within normal result	Between groups	
	Mean (SD)	Paired *t*-test	Mean (SD)	Paired *t*-test	Difference in means	95% CI
CWS-R						
T1 ($n = 111, 1171$)	11.61 (2.90)	–	10.99 (2.91)	–	0.62	1.19 to 0.05, NS
T2 ($n = 111, 1171$)	11.68 (2.89)	–	10.56 (2.60)	–	1.12	1.63 to 0.61, NS
T3 ($n = 111, 1159$)	10.35 (2.65)	–	10.12 (2.49)	–	0.21	0.30 to 0.72, NS
Difference T1–T2	–	–0.298, NS	–	7.537**		
Difference T2–T3	–	6.372**	–	8.633**		
PCQ						
T1 ($n = 110, 1167$)	7.32 (7.66)	–	5.06 (6.71)	–	2.26	3.59 to 0.93, NS
T2 ($n = 110, 1167$)	7.1 (7.44)	–	4.18 (6.19)	–	2.92	4.05 to 1.69*
T3 ($n = 110, 1169$)	4.61 (6.42)	–	3.84 (6.00)	–	0.77	1.95 to 0.41, NS
Difference T1–T2	–	–0.051, NS	–	6.935**		
Difference T2–T3	–	5.752**	–	3.183**		

*$p<0.05$, **$p<0.01$; NS, not significant.
Difference in means calculated by the authors.
Source: adapted from Tyndel *et al.*[101]

TABLE 23 Positive PCQ subscale scores at T2 (1 month after screening) and T3 (6 months after screening)

Outcome	False-positive	Normal mammogram	Mann–Whitney *U*-test	
Positive PCQ at T2			51,561	**
Mean (SD)	13.02 (7.6)	10.81 (6.9)		
Positive PCQ at T3			59,169	NS
Mean (SD)	12.65 (8.9)	11.16 (7.0)		
Benefits of screening more positive at T2			OR	95% CI
No. (%)	112 (55)	1164 (27)	3.17	2.14 to 4.70***
Benefits of screening more positive at T3				
No. (%)	105 (35)	1085 (19)	2.35	1.53 to 3.61***

**$p<0.01$; ***$p<0.001$; NS, not significant.
Source: adapted from Tyndel *et al.*[101]

Overall, the levels of distress were similar in both groups although at T2, on the PCQ, the false-positive group showed greater distress. Additionally, 1 month after the all-clear (T2) those women with false-positive mammograms saw statistically significantly more benefits from screening and were more positive about the benefits of screening at T2 and 6 months later (T3).

Brain *et al.*[102] ($n = 1286$) then used the results of Tyndel *et al.*[101] to investigate which factors, evident at pre-screening, predicted cancer distress at 1 month (T2) and 6 months (T3) after assessment. The results of hierarchical multiple regression showed that, among others, cancer worry at 1 month after screening was predicted by having a false-positive mammogram ($p < 0.05$). This was no longer the case at 6 months after the all-clear. The model accounted for 61% of the variance at T2 and 57% at T3. *Table 24* gives all the variables predictive of cancer distress in the FHBC population.

In order to gain greater understanding of their results and how women valued being in the annual FHBC screening programme, the PIMMS Management Group conducted an interview study of their participants who did ($n = 22$) and did not ($n = 36$) have a false-positive mammogram.[107] As this work has been published only as a poster summary, it is only possible to report an overview of their findings.

The thematic analysis of the interviews by Clements *et al.*[107] ($n = 58$) showed that being part of the FHBC screening programme helped to relieve fear of breast cancer and resulted in women feeling more in control of their family history. Women believed that taking part in screening would enable the earlier detection of cancer and that this would lead to a positive outcome. They also believed that a mammogram was more likely to detect early-stage cancer than actually is the case. They thought that an all-clear result meant that they did not have cancer. Although women who had a false-positive result were initially distressed, when they received the all-clear they had increased feelings of reassurance and security and a greater faith in the screening process than those with an initial all-clear result. They felt that being recalled was positive proof that screening worked.

TABLE 24 Multiple regression showing predictive associations between independent baseline variables and cancer worry at 1 or 6 months

T1 (1 month before screening) variable	T2 (1 month after screening) CWS-R	T3 (6 months after screening) CWS-R
T1 cancer worry	0.54***	0.58***
High perceived lifetime risk of breast cancer	0.09***	0.08**
Relative died of breast cancer in the last year	–	0.05*
Belief in increased risk due to family history	0.09***	0.08***
First attendance at the screening programme	–0.07***	–0.04*
Being recalled for further tests – false-positive	0.06*	–
Low emotion focused coping potential	–0.06*	–0.05*
Use of religion as a coping strategy	0.05**	–
Dispositional optimism	–0.05*	0.00, NS
Low challenge appraisal	–0.04*	–0.02, NS
Substance use for coping	0.04*	–

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; NS, not significant.
Source: adapted from Brain *et al.*[102]

Summary: psychological impact in the family history of breast cancer population

Like the women in the normal-risk population, those in the false-positive FHBC group were significantly more distressed at 1 month after screening than were those with normal mammograms. However, this similarity had disappeared 5 months later, when, although the level of distress was higher than in the normal mammogram group, it was no longer significantly higher in the FHBC false-positive population. The results from the positive PCQ may shed some light on this, as women with a false-positive result scored statistically significantly higher at 1 month after assessment than those with normal mammograms (27%) and reported that they felt more positive about the benefits of screening than at their previous visit (55%). It may be that the women with a FHBC were anticipating a malignant result and felt reassured by having further tests that they remained free of cancer, whereas those in the general population believed they were well and were shocked when they were recalled. A few months later the negative effects of the false-positive experience had diminished for both populations. However, women with a FHBC still reported a statistically significantly more positive attitude to screening at 6 months after their last screening appointment than those who had a normal result in the FHBC population, although multiple regression indicated that a false-positive outcome was predictive of distress 1 month after screening.

Unfortunately, there are currently (2011) no published qualitative interview studies among members of the normal-risk UK population who have experienced a false-positive mammogram to help us better understand these differences and Clements *et al*.'s results.[107] It would seem likely that women in the general population approach breast screening with a different belief; they have no reason to anticipate breast cancer and believe themselves to be well. As Wardle and Pope[120] surmise, their view of themselves as healthy people is challenged by being recalled and the spectre of cancer looms large. This is a very different scenario to those women with a family history of the disease, who may carry the belief that they will one day develop breast cancer.[121] Therefore, rather than seeing two assessments (screening and follow-up tests) as giving increased reassurance of good health, women in the general population may view being recalled as undermining their belief in their good health. In the absence of qualitative interview evidence in the general population, this is, of course, speculative.

Impact of a false-positive mammogram on returning for routine screening

Actual reattendance

The results from the large data registry review of mammography screening attendance in East Anglia by McCann *et al*.[61] ($n = 140{,}387$) show that women who had a false-positive mammogram were less likely than those with normal mammograms to reattend 3 years later for their next routine breast screen (RR 0.97; 95% CI 0.96 to 0.98). Additionally, women who had a biopsy in order to rule out cancer were less likely to reattend than those who were assessed by other methods (RR 0.93; 95% CI 0.89 to 0.97) (*Table 25*).

These results do not replicate those of the earlier registry study, based in central and east London, by O'Sullivan *et al*.[108] ($n = 5649$), which found that, although reattendance rates were lower, there was no statistically significant difference in reattendance levels between women with normal mammograms and those who had false-positive ones. These findings included those who had been put on early recall following further assessment (*Table 26*).

The difference in findings may be due to differences in study design. In McCann *et al*.'s study,[61] participants were invited for their second round of screening, whereas O'Sullivan *et al*.'s participants[108] were being invited for up to their fifth screening round. Therefore, the difference in findings may be partly caused by a general decrease in screening attendance over time.[122] However, this does not explain why those who attended more screenings and therefore had an increased risk of a false-positive mammogram should have similar levels of reattendance to those with normal mammograms in O'Sullivan *et al*.'s study.[108]

TABLE 25 Relative risk of reattendance at second round mammography screening (3 years later) following an invitation

Study group	Reattending (*n*)	Not reattending (*n*)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal mammogram	93,081	15,536	85.7 (85.5 to 85.9)	Baseline
False-positive – all	3981	811	83.1 (82.0 to 84.4)	0.97 (0.96 to 0.98)****
False-positive – no biopsy	3572	706	83.5 (82.4 to 84.6)	0.97 (0.96 to 0.99)****
False-positive – biopsy	409	105	79.6 (76.1 to 83.1)	0.93 (0.89 to 0.97)***

***$p<0.001$; ****$p<0.0001$.

a RR calculated by the authors.

Source: adapted from McCann *et al.*[61]

TABLE 26 Relative risk of reattendance at subsequent rounds of mammography screening (3 years after a false-positive mammogram)

Result at initial screening	Reattending, *n* (%)	Not reattending, *n* (%)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal	3841 (71)	1560 (29)	71 (69.56 to 72.44)	Baseline
False-positive – all	175 (71)	73 (29)	71 (64.28 to 77.72)	0.99 (0.91 to 1.08) NS
False-positive – routine recall	119 (74)	43 (26.5)	74 (66.12 to 81.88)	1.03 (0.94 to 1.13) NS
False-positive – early recall	56 (65)	30 (35)	56 (52.51 to 77.49)	0.92 (0.78 to 1.07) NS

NS, not significant.

a RR calculated by the authors.

Source: adapted from O'Sullivan *et al.*[108]

The study by Brett and Austoker,[59] published in 2001 ($n = 505$), into the psychological consequences of false-positive mammograms also included data on reattendance at the next screening clinic. Their results, which agree with those of McCann *et al.*,[61] showed that women who had a false-positive mammogram were less likely to reattend screening in 3 years' time. It is not clear how many mammography-screening invitations these women had received (*Table 27*).

This study was preceded by the RCT by Meldrum *et al.*,[115] in 1994 ($n = 3083$), which aimed to find out the impact of invitation letters tailored to the outcome of the previous screening round (normal or false-positive) on reattendance. Their results show that, for the standard invitation letter, those with a normal mammogram were more likely to reattend (74%) 3 years later, than those with a false-positive one (70%), although there was no statistically significant difference in the RR of attendance between the two groups (*Table 28*).

However, those with a previous false-positive mammogram who received a tailored invitation were more likely to reattend (82%) than those with a previous normal mammogram (74%). There was a small statistically significant RR for reattendance, RR 1.10 (95% CI 1.00 to 1.21) (*Table 29*).

Very little information is given about the subsequent interviews. Sixty-six women were interviewed about the acceptability and understandability of the tailored letter. No negative comments were made about the letter, but only one person spontaneously mentioned that the letter was tailored to her screening history. Most women appeared not to have paid much attention to the content as they had previously

TABLE 27 Relative risk of reattendance at subsequent rounds of mammography screening (3 years after a false-positive mammogram)

Result at initial screening	Reattending, *n* (%)	Not reattending, *n* (%)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal	120 (92)	10 (08)	92 (87.15 to 96.85)	Baseline
False-positive	319 (85)	56 (15)	85 (81.08 to 88.92)	0.92 (0.86 to 0.98)*

*$p < 0.05$.
a RR calculated by the authors.
Source: adapted from Brett and Austoker.[59]

TABLE 28 Relative risk of receiving a standard invitation letter on reattendance at subsequent rounds of mammography screening (3 years later)

Result at initial screening	Reattending, *n* (%)	Not reattending, *n* (%)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal	583 (74)	208 (26)	74 (71 to 77)	Baseline
False-positive	78 (70)	34 (30)	70 (61 to 78)	0.94 (0.83 to 1.08) NS

NS, not significant.
a RR calculated by the authors.
Source: adapted from Meldrum *et al*.[115]

TABLE 29 Relative risk of receiving a tailored invitation letter on reattendance at subsequent rounds of mammography screening (3 years later)

Result at initial screening	Reattending, *n* (%)	Not reattending, *n* (%)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal	594 (74)	206 (26)	74 (70.47 to 77.53)	Baseline
False-positive	94 (82)	21 (18)	84 (74.23 to 89.77)	1.10 (1.00 to 1.21)*

*$p < 0.05$.
a RR calculated by the authors.
Source: adapted from Meldrum *et al*.[115]

been through the screening process. None of the non-attenders responded that the tailored letter had discouraged them from being screened again.

The final study that reported actual reattendance data is by Orton *et al*.[109] ($n = 1582$). Their findings showed that more women with previous false-positive mammograms reattended mammography screening (92%) than those who had a normal outcome (89%). However, the RR of reattendance was not statistically significant (*Table 30*).

Intended reattendance

The evidence about women's intention to reattend mammography screening when they have had a false-positive mammogram is very limited. It comes only from the final paper from the OPCERG study by Brett and Austoker ($n = 505$).[59] They reported the results from a questionnaire that asked participants which external factors had influenced their decision to attend their next routine screening in 1 months' time. These factors can be seen in *Table 31*, the only items that women reported as worrying influences being magazine or newspaper articles.

TABLE 30 Relative risk of reattendance at subsequent rounds of mammography screening (3 years after a false-positive mammogram)

Result at initial screening	Reattending, *n* (%)	Not reattending, *n* (%)	Reattendance, % (95% CI)	RR[a] (95% CI)
Normal	1362 (89)	170 (11)	89 (87.43 to 90.57)	Baseline
False-positive	46 (92)	4 (08)	92 (84.48 to 99.52)	1.03 (0.95 to 1.13) NS

NS, not significant.
a RR calculated by the authors.
Source: adapted from Orton *et al*.[109]

TABLE 31 External factors influencing attitudes and anxiety about attending the next routine breast screening in women with a previous false-positive mammogram (35 months later)

Item	*n/N*	% (95% CI)	Cause worry %
Magazine or newspaper article	83/288	29 (24 to 34)	11
Television programme	72/288	25 (20 to 30)	9
GP attitude to screening	69/288	24 (19 to 29)	–
Friend	60/288	21 (16 to 26)	–
Poster or leaflet	50/288	17 (13 to 22)	–
Family	47/288	16 (12 to 20)	–
Radio programme	37/288	13 (9 to 17)	–

Source: adapted from Brett and Austoker.[59]

The logistic regression conducted by Brett *et al*.[103] of women with false-positive mammograms 5 months after assessment showed that intention to attend the next screening round was not related to experiencing psychological distress at that time (see *Table 10*).

Summary: impact of false-positive mammograms on reattendance

Figure 8 compares the RRs of the actual reattendance studies. A meta-analysis was not undertaken because of the lack of information about the demographic profiles of participants, making it impossible to assess the studies' homogeneity, together with underlying reservations about the validity of meta-analysis of observational data which can magnify underlying bias and confounding. Two studies showed a statistically significant effect on reattendance [RR 0.97 (95% CI 0.96 to 0.98)][61] and [RR 0.92 (95% CI 0.86 to 0.98)],[59] whereas two did not [RR 0.99 (95% CI 0.91 to 1.08)][108] and [RR 1.03 (95% CI 0.95 to 1.13)].[109] It could be argued that most notice should be taken of the study by McCann *et al*.[61] on the basis of size ($n = 140{,}387$) and quality. However, O'Sullivan *et al*.[108] was also a large study ($n = 5549$) and showed no such effect, although their participants differed from those in the other studies as they were not necessarily on their second screening round and there may be an effect of repeated screening reducing attendance.[122] The smallest study was by Brett and Austoker ($n = 505$)[59] who found that significantly more women with false-positive mammograms failed to reattend the second round of screening, whereas Orton *et al*. ($n = 1582$)[109] found a non-statistically significant trend in the opposite direction.

There is some evidence, from the only RCT in this systematic review, that these reattendance figures can be improved for the false-positive group by sending women tailored invitation letters for the next screening round that refer back to their previous outcome. The data from Brett and Austoker[59] on external influences on reattendance following a false-positive mammogram indicate that the media may have a greater

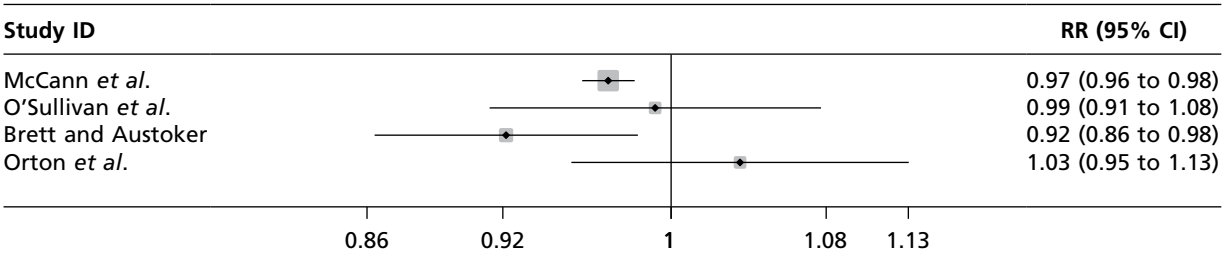

FIGURE 8 Forest plot of the likelihood of failing to reattend the next round of mammography screening following a false-positive mammogram compared with a normal one.

influence than GP, family or friends. The availability of a CNS at screening clinics, as well as assessment clinics, may help to answer questions and concerns resulting from the media.

Overall, it seems likely that the experience of having a false-positive mammogram has a detrimental effect on next-round screening attendance. Future studies should collect demographic information so that their data can be reasonably compared in meta-analysis and this hypothesis can be tested.

Interventions to reduce the impact of false-positive mammograms

As mentioned in *Interventions to reduce the impact of being recalled for further assessment* we found no studies that directly addressed this issue. However, two studies were found that looked at the information needs of women who are recalled after screening. The initial results from the multicentre OPCERG study by Austoker and Ong[112] ($n = 1493$) found that 92% of women were distressed or very distressed when they received their recall letter for assessment following a screening mammogram. Although no standardised measures of anxiety were taken, the quotes from the answers to the open questions in *Table 32* give an indication of the range of responses.

Further results from the questionnaire compared the responses of women who had been given particular items of information in recall letters or leaflets with those who had not. Austoker and Ong[112] found that for all items women wanted more information; this was especially the case when the item was not mentioned in the recall literature. The item that the greatest proportion of women wanted more information about was the reason for their recall (item mentioned in recall literature 36%, item not mentioned in recall literature 46%) (*Table 33*).

Austoker and Ong[112] then compared the level of distress at being recalled with the reported need for more information using an ad hoc questionnaire. They found that women who were distressed or very distressed at being recalled reported a greater need for more information than women who were somewhat or not at all distressed at the prospect of further assessment. The items with the strongest relationship between distress and information need were 'wanting to have more information about why they were recalled', RR 1.84 (95% CI 1.58 to 2.14), and 'more information about how to get more information', RR 1.55 (95% CI 1.27 to 1.89) (*Table 34*).

Austoker and Ong[112] also examined the information in the initial invitation to screening and compared levels of distress, using an ad hoc questionnaire, with whether or not the initial invitation referred to the possibility of recall. They found that when this was not mentioned in the letter or leaflet that women were more likely to be distressed or very distressed if they were recalled, RR 0.76 (95% CI 0.58 to 0.99). They also found an additional benefit from having an information leaflet as well as a recall letter. The RR of finding an aspect of the information about recall reassuring increased statistically significantly when participants were also given an information leaflet, RR 4.04 (95% CI 3.10 to 5.26). Having an additional leaflet also increased women's belief that they understood the assessment procedure, RR 1.27 (95% CI 1.22 to 1.33).

TABLE 32 Reaction of women receiving a recall letter

Reaction	*n* (%) women	Sample comments
Pleased	30 (2)	Very pleased to think I was having a proper check
Neutral/not distressed	87 (6)	I just felt normal
Somewhat distressed	497 (34)	Concerned though not unduly
		I felt rather apprehensive
		Nervous, but I think it is a good thing
		Unpleasantly apprehensive
Distressed	415 (28)	Nervous and very apprehensive
		Anxious and worried
		Frightened and worried
		Worried, afraid
Very distressed	439 (30)	I felt the whole bottom had fallen out of my world
		I felt sick then faint, then I cried then I kept thinking what I have to do if I have cancer
		Worried to death
		Panic stricken, depressed. Convinced I was going to die
		Completely devastated. Reason abandoned me
All women	1468 (100)	

Source: adapted from Austoker and Ong.[112]

TABLE 33 Comparison of the reported need for more information

	Item mentioned in letter/leaflet		Item not mentioned in letter/leaflet	
	Women wanting more information		Women wanting more information	
Item	%	*n/N*	%	*n/N*
Why they were recalled	36	383/1070	46	179/388**
What tests would be done	11	65/606	35	298/847****
Who could come with them	5	44/888	35	148/419****
How to get more information	18	143/783	33	212/633****
Who they would see	13	168/1266	33	62/186****
How long the appointment would take	8	17/222	28	248/900****
How to get to the centre	8	71/854	26	75/290****
How to change the appointment	2	33/1436	–	

**$p<0.01$; ****$p<0.0001$.

Source: adapted from Austoker and Ong.[112]

TABLE 34 Comparison of the level of distress with the need for more information

	Distressed/very distressed women		Somewhat/not distressed women		
	Women wanting more information		Women wanting more information		
Item	%	*n/N*	%	*n/N*	RR (95% CI)
Why they were recalled	48	403/834	26	157/598	1.84 (1.58 to 2.14)****
How to get more information	29	237/811	19	116/616	1.55 (1.28 to 1.89)****
What tests would be done	27	224/828	22	130/599	1.25 (1.03 to 1.51)*
How long the appointment would take	27	173/640	20	93/466	1.35 (1.08 to1.69)**
Who they would see	18	146/826	13	80/598	1.32 (1.03 to 1.70)*
Who could come with them	13	102/762	17	94/557	0.79 (0.61 to 1.03) NS
How to get to the centre	13	83/659	13	64/497	0.98 (0.72 to 1.33) NS
How to change the appointment	2	18/824	3	15/523	0.76 (0.39 to 1.50) NS

*$p<0.05$; **$p<0.01$; ***$p<0.0001$; NS, not significant.
Source: adapted from Austoker and Ong.[112]

Austoker and Ong[112] categorised words and phrases in the recall letter and information sheet, according to responses in the questionnaire and women's additional comments, to produce lists of information that were either reassuring or worrying (*Table 35*).

The study that followed on from this initial research into the information needs of recalled women, by Ong *et al*.,[111] used the categories developed above to evaluate the recall information letters and leaflets sent out by all the assessment centres in the UK. Of the total 87 centres, 84 (97%) sent their materials for evaluation.

The authors found that 99% of the information sent out with the initial screening invitation referred to the possibility of recall (*Table 36*).

When Ong *et al*.[111] compared the list of 'worrying information' with that contained in the recall letters and leaflets they found that 54% of the literature included one of these items (*Table 37*).

An examination of stress-relieving information found that 83% of recall literature contained at least one item (*Table 38*).

The data from the Austoker and Ong[112] study was further analysed by Ong and Austoker[110] ($n = 1493$) to find out women's views of the quality of communication at the assessment clinic, how this related to their level of distress and what role a breast care nurse might play in mitigating this distress. *Table 39* shows that there was a strong link between whether or not participants would have liked to have talked with someone at the assessment centre [RR 1.42 (95% CI 1.15 to 1.74)], whether or not they thought they were given enough information about their physical examination [RR 1.71 (95% CI 1.02 to 2.86)] or X-rays [RR 2.34 (95% CI 1.47 to 3.73)] and whether they were distressed/very distressed or somewhat/not distressed about being recalled.

The results in *Table 40* clearly show that the availability of a breast care nurse at an assessment centre greatly increased the probability that participants would have talked to somebody at the centre about the

TABLE 35 Categorisation of written information in the letters and information leaflets

Reassuring aspects of the information	Worrying aspects of the information
To receive a leaflet describing assessment as well as the recall letter	Receiving the recall letter on a Saturday
To be told that 'being recalled is part of routine (or second stage) screening' or that 'the great majority are found to have normal breasts'	The waiting time between receiving the recall letter and the appointment
To be told in the recall letter that more information can be obtained by phoning the centre	The vagueness of reasons given for recall – for example, 'for a variety of reasons the tests have to be repeated'
To be told that the woman could contact the breast care nurse at the centre	Being told 'not to worry' or 'not to be alarmed'
To be told when women will receive the results	Being told that the mammogram had been 'unclear'
	Being told that the reason for being recalled would be given to the women when they were seen at the centre, the implication being that the centres knew something (assumed to be bad news) that was being withheld from the women
	Using the word 'cancer' specifically in the context of recall in the initial letter of invitation for mammography
	Using the words 'cancer', 'something wrong', 'treatment' or 'abnormality' in the recall letter
	Recall letters stating that 'in the majority of women the results of the second visit show that they do not have cancer' rather than 'most of these (recalled) women are found to have normal breasts'
	Being told in the recall letter/leaflet that assessment was at a hospital rather than a centre/unit/clinic
	Being told that women would be seen 'by a team of specialists'
	Being told that the women could contact a nurse 'counsellor' (rather than a 'breast care nurse', which was reassuring), or that a 'counsellor' would be available at the centre
	Being given detailed descriptions in the recall leaflet about FNA

Source: adapted from Austoker and Ong.[112]

TABLE 36 Probability of information about the possibility of recall being mentioned in the initial screening invitation

	Mentioned in any of the written information:				Mentioned in neither leaflet nor letter nor GP letter, % (no.) of centres
Topic	In the letter, % (no.) of centres	In GP letter, % (no.) of centres	In the leaflet, % (no.) of centres	In both letter and leaflet, % (no.) of centres	
Possibility of recall	46 (39/84)	5 (4/84)	99 (83/84)	45 (38/84)	1 (1/84)

Source: adapted from Ong *et al.*[111]

reason for their recall, RR 0.62 (95% CI 0.59 to 0.66). Where the opportunity to talk to someone was not available women were far more likely to have liked to talk to someone, RR 7.46 (95% CI 4.57 to 12.16). Women were also more likely to have had their assessment tests explained to them and not to need more information about their tests if a breast care nurse was available at the assessment centre (see *Table 40*).

TABLE 37 Particularly worrying information found in the recall letter or leaflet

Topics mentioned	Mentioned in any of the written information:			Mentioned in neither recall leaflet nor recall letter, % (no.) of centres
	In recall letter, % (no.) of centres	In recall leaflet, % (no.) of centres	In both recall leaflet and letter, % (no.) of centres	
One or more worrying items	43 (35/82)	18 (15/82)	7 (6/82)	46 (38/82)
Word 'cancer'	9 (7/82)	10 (8/82)	1 (1/82)	83 (68/82)
Words 'treatment', 'something wrong', 'abnormality', or 'abnormal area of the breast'	20 (16/82)	4 (3/82)	1 (1/82)	78 (64/82)
Word 'hospital'[a]	10 (8/82)	1 (1/82)	0	89 (73/82)
Words 'not to worry'[b]	22 (18/82)	1 (1/82)	0	77 (63/82)
Phrase 'nurse counsellor'	5 (4/82)	9 (7/82)	0	87 (71/82)

a 'Hospital' was only counted when it was mentioned other than in the context of address or directions.

b Similar phrases counted were, 'not to be alarmed', 'not to be concerned', 'not to feel anxious', 'no cause for concern'.

Source: adapted from Ong *et al.*[111]

TABLE 38 Particularly reassuring information in the recall letters or leaflets

Topics mentioned	Mentioned in any of the written information:			Mentioned in neither recall leaflet nor recall letter, % (no.) of centres
	In recall letter, % (no.) of centres	In recall leaflet, % (no.) of centres	In both recall leaflet and letter, % (no.) of centres	
One or more stress-relieving messages	68 (56/82)	33 (27/82)	20 (16/82)	17 (14/82)
Recall is part of second stage/routine screening	46 (77/82)	26 (3/82)	11 (9/82)	38 (31/82)
Most recalled women are found to have normal breasts	28 (23/82)	6 (5/82)	4 (3/82)	30 (25/82)
A substantial number of women are recalled	32 (26/82)	11 (9/82)	1 (1/82)	60 (49/82)

Source: adapted from Ong *et al.*[111]

The final study in this section is the survey of user satisfaction with the Leicestershire breast screening service by Smith *et al.*[113] ($n = 103$). *Table 41* shows that they found that 75% of participants were upset or very upset when they received their recall letter.

The results of the survey about how satisfied the women were with information given in different versions of the recall letter about the reason for their recall and what would happen at the clinic showed that satisfaction varied considerably between the versions of the letter. Women were most satisfied with letter three that gave the most information about what would happen at the clinic and least satisfied with letter one that gave minimal information (*Table 42*).

TABLE 39 Communication at the assessment centre and level of distress

Communication	Distressed/very distressed		Somewhat/not distressed		
	% (*n*/*N*)	95% CI	% (*n*/*N*)	95% CI	RR (95% CI)
Women who had not had the opportunity to talk with a health worker at the centre about the reason for recall	33 (275/835)	0.30 to 0.36	32 (191/597)	0.28 to 0.36	1.03 (0.88 to 1.20) NS
Women who would have liked to talk about the reason for recall	26 (214/835)	0.23 to 0.29	18 (108/597)	0.15 to 0.21	1.42 (1.15 to 1.74)****
Women who thought they were not given enough information about their physical examination	6 (46/757)	0.04 to 0.08	4 (20/563)	0.02 to 0.05	1.71 (1.02 to 2.86)****
Women who thought they were not given enough information about their X-rays	9 (72/773)	0.07 to 0.12	4 (22/553)	0.03 to 0.06	2.34 (1.47 to 3.73)****

****$p<0.0001$; NS, not significant.

a RR calculated by the authors.

Source: adapted from Ong and Austoker.[110]

TABLE 40 Communication at the assessment centre and the role of breast care nurses

Communication	Centres where women were not systematically provided with the opportunity to talk immediately before tests		Centres where the breast care nurse provided women with the opportunity to talk in private immediately before tests		
	% (*n*/*N*)	95% CI	% (*n*/*N*)	95% CI	RR (95% CI)
Women who had talked at the centre about the reason for recall					
With 'somebody at the centre'	58 (611/1055)	0.55 to 0.61	93 (374/401)	0.90 to 0.96	0.62 (0.59 to 0.66)****
With a doctor or radiologist	31 (323/1035)	0.28 to 0.34	7 (26/391)	0.04 to 0.10	4.69 (3.20 to 6.88)****
With a nurse	9 (97/1035)	0.08 to 0.11	60 (234/391)	0.55 to 0.65	0.16 (0.13 to 0.19)****
Women who would have liked to talk about reason for recall	30 (310/1039)	0.27 to 0.33	4 (16/400)	0.02 to 0.06	7.46 (4.57 to 12.16)****
Women who stated that the tests they had were not explained to them					
Physical examination by a doctor	8 (82/981)	0.07 to 0.10	2 (7/381)	0.01 to 0.04	4.55 (2.12 to 9.75)****
X-rays	9 (88/996)	0.05 to 0.11	1 (5/379)	0.00 to 0.03	6.70 (2.74 to 16.36)****
Ultrasound	9 (39/413)	0.07 to 0.13	2 (5/212)	0.01 to 0.05	4.00 (1.60 to 10.01)***
Women who wanted more information about their tests					
Physical examination by a doctor	6 (59/964)	0.05 to 0.08	2 (7/378)	0.01 to 0.04	3.31 (1.52 to 7.17)***
X-rays	7 (68/971)	0.05 to 0.09	2 (8/376)	0.01 to 0.04	3.29 (1.60 to 6.78)***
Ultrasound	10 (39/401)	0.07 to 0.13	3 (6/209)	0.01 to 0.06	3.39 (1.46 to 7.87)***

***$p<0.001$; ****$p<0.0001$.

a RR calculated by the authors.

Source: adapted from Ong and Austoker.[110]

TABLE 41 How women felt when they received their invitation letter to return for further assessment

Reaction	*n*/%
Positive (e.g. 'glad to be in such capable hands')	4
Neutral (e.g. 'I wasn't bothered')	10
Surprised	11
Upset (e.g. 'anxious', 'worried', 'upset')	44
Very upset (e.g. 'terrified', 'extremely anxious')	31
Total	100

Not all participants answered all questions.
Source: adapted from Smith *et al.*[113]

TABLE 42 Satisfaction of women with information about why they had to return to clinic and what would happen there

Letter version	Reasons for recall *n* (%)	Events at the clinic *n* (%)
1. Minimal information about possible tests	15(50)	17 (63)
2. Offer of telephoning a breast care nurse for more information	25 (71)	24 (74)
3. Similar to letter one but with more detail about what would happen at the clinic	26 (81)	27 (90)
All versions	66 (68)	68 (76)
Chi-squared	7.243	5.817
p-value	0.027	0.055

Source: adapted from Smith *et al.*[113]

When participants were asked if they would have telephoned a breast care nurse if one was available, 98% said they would and 100% of the women who received letter two with the breast care nurse's telephone number used this facility (*Table 43*).

Summary: interventions to reduce the impact of false-positive mammograms

The results from these studies evidently show that most women want clear information about the reason for their recall and what will happen to them at their assessment and that having this information can reduce stress levels; however, too much information about the process of FNA was considered stressful. There was a benefit from having a recall leaflet as well as a letter, as this increased the likelihood that women believed they understood what would happen to them. The level of stress experienced by women when they were recalled was increased if the possibility of recall was not mentioned in the original screening invitation. The value of having a breast care nurse to talk to is very apparent by increasing satisfaction with the process of assessment, reducing distress and satisfying the need to have someone at the clinic to talk to. It is also interesting to note the effect of language in the recall literature, with 54% of the recall literature containing at least one worrying item and 83% containing at least one stress-relieving message.

TABLE 43 Likelihood of phoning a breast care nurse

Letter version	Answer	*n*
1 – would telephone the BCN	Yes	25
	No	0
2 – did telephone the BCN	Yes	13
	No	0
3 – would telephone the BCN	Yes	17
	No	1

BCN, breast care nurse.
Source: adapted from Smith *et al*.[113]

Chapter 4 Discussion

Statement of principal findings

The aim of this systematic review was to identify the psychological impact on women of false-positive screening mammograms and any evidence for the effectiveness of interventions designed to reduce this impact. We were also looking for evidence of effects in subgroups of women.

Our searches retrieved 4423 titles and abstracts after deduplication. When screening was complete, we found five systematic reviews,[62,82,85–87] one meta-analysis and 11 primary studies that met our inclusion criteria. No studies were found that were either about or reported subgroups of women from different ethnic, socioeconomic or other groups within the general screening population. One study[101,102,107] was found that included women who had a false-positive mammogram and a FHBC. The quality of the primary research was variable: we found one poor-quality RCT[115] and 10 observational studies.[55,59,61,101–114] The best-quality observational research was conducted by OPCERG and PIMMS groups and the observational studies reporting reattendance rates. However, even here there were shortcomings in reporting key information. Overall, the main weaknesses in reporting were a failure to consider the possible effects of bias and confounding on the results and a failure to report participants' demographic and other characteristics, making the interpretation of the results very difficult. These quality indicators appear to have been overlooked, as, in most cases, there was no consideration of the limitations of the methods or conduct of the study. Therefore, the results of this systematic review must be treated with caution, not just because they come from a limited number of observational studies, but also because many of these studies lack methodological robustness.

General population

The studies of the psychological impact of false-positive mammograms in the general population gave conflicting results. When disease-specific measures were used (PCQ), an enduring negative impact was found that lasted until 35 months from the last assessment. The degree of distress found was related to the level of invasiveness of the method of assessment used so that, at 35 months, women who had a biopsy were more distressed (RR 2.07; 95% CI 1.22 to 3.52) than women who had FNA (RR 1.80; 95% CI 1.17 to 2.77) and, non-significantly, further mammography (RR 1.28; 95% CI 0.82 to 2.00). Additionally, women placed on early recall were also at a greater RR of distress (RR 1.82; 95% CI 1.22 to 2.72). The greatest RR of distress was felt at 5 months after assessment and was significant for all assessment procedures.

Further evidence of psychological distress came from results of a comparison of the frequency of breast self-examination of women who had false-positive or normal mammograms. This found that significantly more women who were recalled examined themselves once a week or more often, which may be taken as a proxy for anxiety although the meaning of this behaviour is unclear.

Conversely, when generic measures of clinical levels of general anxiety and depression were used (HADS and GHQ-28) no significant differences were found between the two groups at 6 weeks after assessment and 3 months after screening.

Therefore, it may be reasonable to speculate that, for those in the general population, having a false-positive screening mammogram can cause breast cancer-specific psychological distress that may endure for up to 3 years. However, it is less likely that there will be general anxiety detectable at clinically recognisable levels.

Family history of breast cancer population

These results draw a slightly different picture to those in the general-risk population. Here the psychological distress of the false-positive group was statistically significantly greater than the normal group only at 1 month after screening [negative PCQ, difference in means 2.92 (95% CI 4.05 to 1.69)]. At the same time the false-positive group also scored significantly higher on the positive PCQ than those with normal mammograms (Mann–Whitney *U*-test 51,561; $p < 0.05$). They also rated the benefits of screening more highly than those with normal mammograms at 1 (T2) and 6 months (T3) after screening on an ad hoc questionnaire [OR: T2, 3.17 (95% CI 2.14 to 4.70); T3, 2.35 (95% CI 1.53 to 3.61)]. These results may appear to be conflicting at first glance. However, the summary results from the unpublished interview study suggest that the women in the false-positive group may have been rationalising their anxiety at being recalled by reassuring themselves that this meant that the programme was thorough and would detect early cancer that could be treated.

Impact of a false-positive mammogram on returning for routine screening

The evidence for the impact of having a false-positive mammogram on returning for the next screening round is conflicting. It comes mainly from four retrospective observational studies[55,61,108,109] that collected data from registries and other NHS databases. The weight of evidence, in terms of the numbers of participants, is that women with false-positive mammograms are less likely to return for their next round of screening than women with normal mammograms. The largest study[61] with this finding ($n = 140,387$) had a RR of not returning of 0.97 (95% CI 0.96 to 0.98). Two studies[108,109] with a combined population of 7231 found that there was no such association. Evidence from a poor-quality RCT[115] suggests that this finding can be reversed if women are given screening invitation letters that are tailored to the outcome of their last screening (RR 1.10; 95% CI 1.00 to 1.21).

Interventions to reduce the impact of false-positive mammograms

The above evidence suggests that in the general population there is a negative psychological impact from having a false-positive mammogram that may endure for 3 years and may deter women from attending their next round of screening. Unfortunately, we were unable to find any studies that directly addressed these problems. Nevertheless, we identified two studies[110–113] that looked at the information needs of women who were recalled and the importance of communication. These studies showed that women wanted clear information about the reason they had been recalled. However, the ability of clinicians to address this is limited by the need to stage the information to ensure direct and face-to-face support in cases where radiologists are reasonably sure that the screening mammogram indicates cancer. Women also indicated that they wanted clear information about what would happen at their assessment, as well as access to a breast care nurse or CNS to talk through their concerns. Satisfaction with the service increased if women were sent a recall leaflet as well as a letter as participants believed that they had better understanding of what would happen to them at the assessment clinic. The importance of the language used in the recall literature was also evident, with particular words and phrases reducing or increasing stress. The research by the OPCERG was used to produce national guidelines on improving the quality of written information sent to women who were recalled for assessment in 1998.[114]

These intervention studies are more than 10 years old, and it is unknown whether or not the recommendations in the national guidelines have been implemented. There is currently no national recall information leaflet similar to the NHS breast-screening leaflet.

Comparison with other systematic reviews

With the exception of the systematic review by Bankhead *et al.*,[84] the quality of the included secondary research was not very high. Of particular concern was the lack of consideration of the effects of bias and confounding in their included observational studies.

Our results agree with those of previous systematic reviews and meta-analyses. In particular, we agree with all these evidence syntheses that there can be negative psychological consequences from having a false-positive mammogram. We also supported their finding that having a false-positive mammogram increased the frequency of breast self-examination, which may be a proxy for anxiety. The meta-analysis by Salz *et al*.[85] that compared outcomes measured by disease-specific or generic measures of psychological distress agrees with our finding that the type of outcome measure used can affect whether or not an outcome is found to be statistically significant, although, unlike Salz *et al*.,[85] we were unable to find evidence of anxiety at clinical levels. We also agree with the finding that the evidence about reattendance is conflicted. However, the weight of evidence suggests that women with this result are less likely to reattend.

Strengths and limitations

The strengths of this systematic review are that it was conducted by an independent research team using robust methods. Our searches were comprehensive and we believe that we have retrieved all includable studies.

The systematic review may have been influenced by publication bias. However, there were insufficient studies in each domain to produce a funnel plot that would give meaningful information.

The robustness of the findings of this systematic review is limited by the reliability of the included studies. With the exception of one weak RCT all the studies were observational and so subject to the risks of bias and confounding that are associated with these designs, compounded by lack of reporting key information such as the baseline characteristics of the groups. However, the nature of the subject under study necessitates observational designs; therefore, great care should be taken in matching groups and reporting their characteristics.

Chapter 5 Conclusions

We conclude that the experience of having a false-positive screening mammogram, in the general risk of breast cancer population, can cause breast cancer-specific psychological distress that may endure for up to 3 years. However, it is less likely that there will be general anxiety detectable at clinically recognisable levels. The likelihood of women experiencing distress may be determined by the degree of invasiveness of the assessment procedure, with more invasive techniques increasing the probability of psychological distress.

The strongest evidence suggests that the distress caused by a false-positive mammogram may be sufficient to deter an additional 3% of women from attending their next breast cancer screening appointment.

It is important to provide women who are recalled with clear, carefully worded information about the reason for their recall and the process of the assessment (but not in such detail that they become distressed without the support of the screening staff being present), and to have a breast care nurse or CNS available to answer any concerns they may have.

There is some evidence that having a subsequent round of screening invitation letter that refers to the outcome of the previous screening round may encourage women with false-positive mammograms to reattend.

For women with a FHBC, having a false-positive mammogram, while increasing levels of distress, may also reassure them that they are being well looked after and that early cancer will be detected that could be treated.

Additionally, the evidence suggests that women with false-positive mammograms are at three times greater risk of interval cancer than those with normal mammograms and are more than twice as likely to have cancer detected at the next screening round. The reasons for this finding need further explanation.

Implications for health care

The evidence suggests that the availability of a CNS at screening clinics as well as assessment clinics may help to allay concerns raised by information sources other than official literature. There is also some evidence that including a reference to the previous recall in the next round invitation letter may facilitate attendance.

Research recommendations

The evidence found by this systematic review in the general population is at least 10 years old. Up-to-date studies are needed that reflect current screening practice.

1. A qualitative interview study of the general population of women who have had false-positive screening mammograms, in order to understand what this experience means to them.
2. Well-designed observational studies, in the general screening population, that use disease-specific and generic outcome measures in order to determine the level of severity of negative psychological outcomes. Including studies of women from different ethnic and socioeconomic groups.
3. The routine collection of demographic information in observational studies so that future systematic reviews may be able to judge whether or not the pooling of data is possible.

4. Currently there is no standard national recall letter following a suspicious screening mammogram. There should be a national survey of the recall literature sent out from NHSBSP services to see if the national guidelines produced in 1998 are being adhered to, followed by the development of such a letter.
5. There is some evidence to suggest that there may be a relationship between tailored invitation letters for next screening round for women who have had false-positive mammograms and reattendance. A well-designed RCT would be able to help us understand whether or not this relationship exists and a nested qualitative study would give insight into the important features of such a letter.
6. Developmental and pilot work of interventions both to relieve the distress of false-positive mammograms and to encourage women with this outcome to reattend routine screening. Promising interventions should then be tested in well-designed RCTs sufficiently powered to allow for subgroup analysis.

A list of ongoing studies can be found in *Appendix 4*. We found no ongoing UK studies about false-positive mammograms.

Acknowledgements

We would like to acknowledge Jenny Lowe and Sue Whiffin for their administrative support and especially Jaime Peters for her statistical advice.

Expert Advisory Group

We would particularly like to thank our expert advisors:

Ms Kate Blackmore	Patient representative.
Dr Russell Davies	Consultant radiologist, Royal Devon & Exeter NHS Foundation Trust.
Professor Jenny Hewison	Professor of the Psychology of Healthcare, University of Leeds.
Ms Sue Milward	Patient representative.
Professor Carl Roobottom	Consultant radiologist, Plymouth Hospitals NHS Trust.
Dr Jim Steel	Director of Screening, Consultant breast radiologist, Plymouth Hospitals NHS Trust.

Contributions of authors

Mary Bond (Senior Research Fellow, HTA)	Provided project management, wrote the protocol, codesigned the searches, led the systematic review, wrote and edited the report. This systematic review is part of a continuing research project and PhD supported by the Collaboration for Leadership in Applied Health Research and Care (PenCLAHRC).
Toby Pavey (Research Fellow, HTA)	Acted as second reviewer and commented on the draught report.
Karen Welch (Information Specialist)	Designed and undertook literature searches.
Chris Cooper (Information Specialist)	Designed and undertook literature searches and commented on the draught report.
Ruth Garside (Senior Research Fellow, HTA)	Qualitative research advisor and commented on the draught report.
Sarah Dean (Senior Lecturer, HSR)	Health psychology advisor and commented on the draught report.
Chris Hyde (Professor Public Health and Clinical Epidemiology)	Commented on the draught report and was overall director and guarantor of the report.

About the Peninsula Technology Assessment Group

The Peninsula Technology Assessment Group (PenTAG) is part of the Institute of Health Service Research at the Peninsula Medical School. PenTAG was established in 2000 and currently has three major work streams; independent Health Technology Assessments for NICE and the NIHR HTA Programme, systematic reviews as part of the Cochrane Collaboration Heart Group and evidence synthesis work in relation to the needs of PenCLAHRC, as well as for other local and national decision-makers.

The group is multidisciplinary and draws on individuals' backgrounds in public health, health services research, computing and decision analysis, systematic reviewing, statistics and health economics. The Peninsula Medical School is a school within the Universities of Plymouth and Exeter. The Institute of Health Research is made up of discrete but methodologically related research groups, among which Health Technology Assessment is a strong and recurring theme. Recent PenTAG HTA projects include:

Bond M, Rogers G, Peters J, Anderson R, Hoyle M, Miners A, *et al*. The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of TA111): a systematic review and economic model. *Health Technol Assess* 2012;**16**(21).

Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, *et al*. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2010;**14**(2).

Hoyle M, Crathorne L, Garside R, Hyde C. Ofatumumab (Arzerra®) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab: a critique of the submission from GSK. *Health Technol Assess* 2011;**15**(Suppl. 1).

Rogers G, Hoyle M, Thompson-Coon J, Pitt M, Moxham T, Liu Z, *et al*. Dasatinib and nilotinib for imatinib-resistant or intolerant chronic myeloid leukaemia: a systematic review and economic evaluation. *Health Technol Assess* 2012;**16**(22).

References

1. Forrest P. *Breast cancer screening. Report to the Health Ministers of England, Wales, Scotland and Northern Ireland*. London: HMSO; 1986.
2. The NHS Information Centre PHIT. *Breast screening results from the NHSBSP 2009–10*. 2011. URL: www.cancerscreening.nhs.uk/breastscreen/statistics.html (cited 28 March 2011).
3. World Health Organisation (WHO) International Agency for Research on Cancer. *Breast cancer screening. International Agency for Research on Cancer*. WHO; 2002.
4. Cancer Research UK. *Breast cancer-UK incidence statistics*. 2010. URL: http://info.cancerresearchuk.org (cited 4 May 2010).
5. Beral V, Banks E, Reeves G, Wallis M. Hormone replacement therapy and high incidence of breast cancer between mammographic screens. *Lancet* 1997;**349**:1103–4. http://dx.doi.org/10.1016/S0140-6736(05)62328-8
6. Quinn MJ, Cooper N, Rachet B, Mitry E, Coleman MP. *Survival from cancer of the breast in women in England and Wales up to* 2001. *Br J Cancer* 2008;**99**(Suppl. 1):S53–S55. http://dx.doi.org/10.1038/sj.bjc.6604587
7. Parsa P, Parsa B. Effects of reproductive factors on risk of breast cancer: a literature review. *Asian Pac J Cancer Prev* 2009;**10**:545–50.
8. Office for National Statistics. *Cancer statistics registrations: registrations of cancer diagnosed in 2007, England*. London: National Statistics; 2010.
9. ISD Scotland. *ISD Cancer Information Programme*. 2009. URL: www.isdscotland.org (cited 4 November 2012).
10. Welsh Cancer Intelligence and Surveillance Unit. *Cancer incidence in Wales*. 2010. URL: www.wales.nhs.uk (cited 4 November 2010).
11. Office for National Statistics. *Breast Cancer: incidence rates rise, mortality rates fall*. 2011. URL: www.statistics.gov.uk/CCI/nugget.asp?ID=575 (cited 29 March 2011).
12. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;**343**.
13. U.S. Preventative Services Task Force. Addendum and correction: screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2010;**152**:688.
14. Bleyer A. US breast cancer mortality is consistent with European data. *BMJ* 2011;**343**.
15. de Koning H. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;**343**:d4411. http://dx.doi.org/10.1136/bmj.d4411
16. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011;**1**:CD001877.
17. Shapiro S, Venet W, Strax P. Selection, follow-up, and analysis in the health insurance plan study: a randomized trial with breast cancer screening. *J Natl Cancer Inst Monogr* 1985;**67**:65–74.
18. Strax P. Mass screening for breast cancer. *Rev Fr Gynecol Obstet* 1973;**68**:457–66.

19. Fink R, Shapiro S, Roester R. Impact of efforts to increase participation in repetitive screenings for early breast cancer detection. *Am J Public Health* 1972;**62**:328–36. http://dx.doi.org/10.2105/AJPH.62.3.328

20. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, *et al.* Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *BMJ* 1988;**297**:943–8. http://dx.doi.org/10.1136/bmj.297.6654.943

21. Andersson I. Radiographic screening for breast carcinoma. II. Prognostic considerations on the basis of a short-term follow-up. *Acta Radiol Diagn* 1981;**22**:227–33. http://dx.doi.org/10.1111/j.1748-1716.1981.tb06887.x

22. Andersson I. Radiographic screening for breast carcinoma. I. Program and primary findings in 45–69 year old women. *Acta Radiol Diagn* 1981;**22**:185–94.

23. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Scand J Caring Sci* 2002;**359**:909–19.

24. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, *et al.* Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health & Welfare. *Lancet* 1985;**i**:829–32.

25. Tabar L, Gad A. Screening for breast cancer: the Swedish trial. *Radiology* 1981;**138**:219–22.

26. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radio Clin North Am* 1992;**30**:187–210.

27. Alexander F, Roberts MM, Lutz W, Hepburn W. Randomisation by cluster and the problem of social class bias. *J Epidemiol Community Health* 1989;**43**:29–36. http://dx.doi.org/10.1136/jech.43.1.29

28. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, *et al.* The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. *Br J Cancer* 1994;**70**:542–8. http://dx.doi.org/10.1038/bjc.1994.342

29. Alexander FE, Anderson TJ, Brown HK, Forrest APM, Hepburn W, Kirkpatrick AE, *et al.* 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;**353**:1903–8. http://dx.doi.org/10.1016/S0140-6736(98)07413-3

30. Baines CJ. The Canadian National Breast Screening Study: a perspective on criticisms. *Ann Intern Med* 1994;**120**:326–34.

31. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study – 2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;**92**:1490–9. http://dx.doi.org/10.1093/jnci/92.18.1490

32. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study – 1: breast cancer mortality after 11 to 16 years of follow-up. *Ann Intern Med* 2002;**137**:E-305.

33. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study: update on breast cancer mortality. *J Natl Cancer Inst Monogr* 1997;**22**:37–41.

34. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Follow up after 11 years – update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat* 1997;**45**:263–70. http://dx.doi.org/10.1023/A:1005872617944

35. Frisell J, Glas U, Hellstrom L, Somell A. Randomized mammographic screening for breast cancer in Stockholm. Design, first round results and comparisons. *Breast Cancer Res Treat* 1986;**8**:45–54. http://dx.doi.org/10.1007/BF01805924

36. Frisell J, Eklund G, Hellstrom L, Lidbrink E, Rutqvist LE, Somell A. Randomized study of mammography screening – preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat* 1991;**18**:49–56. http://dx.doi.org/10.1007/BF01975443

37. Frisell J, Eklund G, Hellstrom L, Glas U, Somell A. The Stockholm breast cancer screening trial – 5-year results and stage at discovery. *Breast Cancer Res Treat* 1989;**13**:79–87. http://dx.doi.org/10.1007/BF01806553

38. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, *et al.* The Gothenburg Breast Screening Trial. *Cancer* 2003;**97**:2387–96. http://dx.doi.org/10.1002/cncr.11361

39. Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Eriksson O, *et al.* The Gothenburg breast screening trial. *Cancer* 1997;**80**:2091–9.

40. UK Trial. First results on mortality reduction in the UK Trial of Early Detection of Breast Cancer. UK Trial of Early Detection of Breast Cancer Group. *Lancet* 1988;[0140-6736 (Print)].

41. Moss S. A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group. *J Med Screen* 1999;**6**:144–8.

42. Moss S, Waller M, Anderson TJ, Cuckle H (writing committee for the Trial Management Group). Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures. *Br J Cancer* 2005;**92**:955–60. http://dx.doi.org/10.1038/sj.bjc.6602395

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

44. Moss S, Thomas I, Evans A, Thomas B, Johns L. Randomised controlled trial of mammographic screening in women from age 40: results of screening in the first 10 years. *Br J Cancer* 2005;**92**:949–54. http://dx.doi.org/10.1038/sj.bjc.6602396

45. Kopans DB, Smith RA, Duffy SW. Mammographic Screening and Overdiagnosis. *Radiology* 2011;**260**:616–20. http://dx.doi.org/10.1148/radiol.11110716

46. Otto SJ, Fracheboud J, Looman C, Broeders M, Boer R, Hendriks J, *et al.* Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;**361**:1411–17. http://dx.doi.org/10.1016/S0140-6736(03)13132-7

47. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen THH, *et al.* Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;**17**:25–30. http://dx.doi.org/10.1258/jms.2009.009094

48. Wald NJ, Law MR, Duffy SW. Breast screening saves lives. *BMJ* 2009;**339**.

49. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, *et al.* Effect of screening and adjuvant therapy on mortality from breast cancer. 2005;**353**:1784–92.

50. Duffy S. *Re: Estimate of breast screening benefit was 6 times too large*. 2009. URL: www.bmj.com/cgi/eletters/338/jan27_2/b86 (cited 11 October 2011).

51. Marteau TM, Cook R, Kidd J, Michie S, Johnston M, Slack J, *et al.* The psychological effects of false-positive results in prenatal screening for fetal abnormality: a prospective study. *Prenat Diagn* 1992;**12**:205–14. http://dx.doi.org/10.1002/pd.1970120309

52. French DP, Maissi E, Marteau TM. The psychological costs of inadequate cervical smear test results: three-month follow-up. *Psychooncology* 2006;**15**:498–508. http://dx.doi.org/10.1002/pon.980

53. Lampic C, Thurfjell E, Bergh J, Sjoden P-O. Short- and long-term anxiety and depression in women recalled after breast cancer screening. *Eur J Cancer* 2001;**37**:463–9. http://dx.doi.org/10.1016/S0959-8049(00)00426-3

54. Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. *J Epidemiol Community Health* 1997;**51**:705–10. http://dx.doi.org/10.1136/jech.51.6.705

55. Sutton S, Saidi G, Bickler G, Hunter J. Does routine screening for breast cancer raise anxiety? Results from a three wave prospective study in England. *J Epidemiol Community Health* 1995;**49**:413–18. http://dx.doi.org/10.1136/jech.49.4.413

56. Aro AR, Pilvikki Absetz S, van Elderen TM, van der Ploeg E, van der Kamp LJT. False-positive findings in mammography screening induces short-term distress – breast cancer-specific concern prevails longer. *Eur J Cancer* 2000;**36**:1089–97. http://dx.doi.org/10.1016/S0959-8049(00)00065-4

57. Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. *Br J Cancer* 1990;**62**:1018–22. http://dx.doi.org/10.1038/bjc.1990.430

58. Burman ML, Taplin SH, Herta DF, Elmore JG. Effect of false-positive mammograms on interval breast cancer screening in a health maintenance organization. *Ann Intern Med* 1999;**131**:1–6.

59. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. *J Public Health* 2001;**23**:292–300. http://dx.doi.org/10.1093/pubmed/23.4.292

60. Lampic C, Thurfjell E, Sjoden P-O. The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. *Eur J Cancer* 2003;**39**:1730–7. http://dx.doi.org/10.1016/S0959-8049(02)00451-3

61. McCann J, Stockton D, Godward S. Impact of false-positive mammography on subsequent screening attendance and risk of cancer. *Breast Cancer Res* 2002;**4**:R11. http://dx.doi.org/10.1186/bcr455

62. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007;**146**:502–10.

63. NHS Breast Screening Programme. *Quality assurance guidelines for breast cancer screening radiology.* Wilson R, Liston J, editors. 59. 2011. Sheffield: NHS Cancer Screening Programmes; 2011.

64. NHS Breast Screening Programme. *Clinical guidelines for breast cancer screening assessment*. Liston J, Wilson R, editors. 49. 2010. Sheffield: NHS Cancer Screening Programmes; 2011.

65. NHS Cancer Screening Programme. *Quality assurance guidelines for surgeons in breast cancer screening.* Sibbering M, Watkins R, Winstanely J, Patnick J, editors. 20. 2009. Sheffield: NHS Cancer Screening Programme; 2011.

66. NHS Breast Screening Programme. *Quality assurance guidelines for clinical nurse specialists in breast cancer screening*. 29. 2008. Sheffield: NHS Cancer Screening Programmes; 2011.

67. Cockburn J, De Luise T, Hurley S, Clover K. Development and validation of the PCQ: a questionnaire to measure the psychological consequences of screening mammography. *So Sci Med* 1992;**34**:1129–34. http://dx.doi.org/10.1016/0277-9536(92)90286-Y

68. Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. *Ann Intern Med* 1991;**114**:657–61.

69. Lerman C, Trock B, Rimer BK, Jepson C, Brody D, Boyce A. Psychological side effects of breast cancer screening. *Health Psychol* 1991;**10**:259–67. http://dx.doi.org/10.1037/0278-6133.10.4.259

70. Watson M, Lloyd S, Davidson J, Meyer L, Eeles R, Ebbs S, *et al.* The impact of genetic counselling on risk perception and mental health in women with a family history of breast cancer. *Br J Cancer* 1999;**79**:868–74. http://dx.doi.org/10.1038/sj.bjc.6690139

71. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. *Int J Behav Med* 1997;**4**:92–100. http://dx.doi.org/10.1207/s15327558ijbm0401_6

72. Goldberg D. *The detection of psychiatric illness by questionnaire*. Oxford: Oxford University Press; 1972.

73. Goldberg D, Williams P. A User's guide to the General Health Questionnaire. London: NFER-Nelson; 1988.

74. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983;**67**:361–70. http://dx.doi.org/10.1111/j.1600-0447.1983.tb09716.x

75. Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *J Pers Soc Psychol* 1994;**67**:1063–78. http://dx.doi.org/10.1037/0022-3514.67.6.1063

76. Speilberger C. *Manual for the State–Trait Anxiety Inventory: STAI (form Y).* Palo Alto, CA: Consulting Psychologists Press; 1983.

77. NHS Centre for Reviews and Dissemination. *Systematic reviews: CRD's guidance for undertaking reviews in health care*. York: NHS Centre for Reviews and Dissemination; 2009.

78. Egger M, Smith GD, Altman DG. *Systematic reviews in health care: meta-analysis in context*. 2nd edn. London: BMJ; 2001. http://dx.doi.org/10.1002/9780470693926

79. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**.

80. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;**357**:1191–4. http://dx.doi.org/10.1016/S0140-6736(00)04337-3

81. Elm Ev, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8. http://dx.doi.org/10.1136/bmj.39335.541782.AD

82. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. *Psycho Oncol* 2005;**14**:917–38. http://dx.doi.org/10.1002/pon.904

83. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005;**33**:1064–5. http://dx.doi.org/10.1136/bmj.38636.593461.68

84. Bankhead CR, Brett J, Bukach C, Webster P, Stewart-Brown S, Munafo M, *et al.* The impact of screening on future health-promoting behaviours and health beliefs: a systematic review. *Health Technol Assess* 2003;**7**(42).

85. Salz T, Richman AR, Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes. *Psycho Oncol* 2010;**19**:1026–34. http://dx.doi.org/10.1002/pon.1676

86. Hafslund B, Nortvedt MW. Mammography screening from the perspective of quality of life: a review of the literature. *Scand J Caring Sci* 2009;**23**:539–48. http://dx.doi.org/10.1111/j.1471-6712.2008.00634.x

87. Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE. Clinical guidelines. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Ann Intern Med* 2007;**146**:516.

88. Meystre-Agustoni G, Paccaud F, Jeannin A, Dubois-Arber F. Anxiety in a cohort of Swiss women participating in a mammographic screening programme. *J Med Screen* 2001;**8**:213–19. http://dx.doi.org/10.1136/jms.8.4.213

89. Beck AT, Erbaugh J, Ward CH, Mock J, Mendelsohn M. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;**4**:561–71. http://dx.doi.org/10.1001/archpsyc.1961.01710120031004

90. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults - evaluation of a short-form of the CES-D. *Am J Prev Med* 1994;**10**:77–84.

91. Boer H. *Psychological effects of breast cancer screening*. Twente: University of Twente; 1993.

92. Goldberg DP. *Manual of the General Health Questionnaire*. Windsor: NFER-Nelson Publishing; 1978.

93. Walker LG, Cordiner CM, Gilbert FJ, Needham G, Deans HE, Affleck IR, *et al.* How distressing is attendance for routine breast screening? *Psycho Oncol* 1994;**3**:299–304. http://dx.doi.org/10.1002/pon.2960030406

94. Derogatis L, Lipman RS, Rickels K, Uhlenhut E, Covi L. Hopkins Symptom Checklist (HSCL) – a self-report symptom inventory. *Behav Sci* 1974;**19**:1–15. http://dx.doi.org/10.1002/bs.3830190102

95. Kellner R. *Abridged manual of the illness attitude scales*. Albuquerque, NM: University of New Mexico; 1987.

96. Horowitz M, Wilner N, Alvarez W. Impact of event scale – measure of subjective stress. *Psychosom Med* 1979;**4**:209–18.

97. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, *et al.* Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002;**32**:959–76. http://dx.doi.org/10.1017/S0033291702006074

98. McNaire D, Lorr M, Droppleman LF. *Manual for the profiles of moods scale*. San Diego, CA: Educational and Industrial Testing Services; 1971.

99. Rakowski W, Andersen M, Stoddard AM, Urban N, Rimer BK, Lane DS, *et al.* Confirmatory analysis of opinions regarding the pros and cons of mammography. *Health Psychology* 1997;**16**:433–42. http://dx.doi.org/10.1037/0278-6133.16.5.433

100. Derogatis L. *Administration, scoring and procedures manual 1 to the (revised) version*. Baltimore, MD: Johns Hopkins University School of Medicine; 1977.

101. Tyndel S, Austoker J, Henderson BJ, Brain K, Bankhead C, Clements A, *et al.* What is the psychological impact of mammographic screening on younger women with a family history of breast cancer? Findings from a prospective cohort study by the PIMMS Management Group. *J Clin Oncol* 2007;**25**:3823–30. http://dx.doi.org/10.1200/JCO.2007.11.0437

102. Brain K, Henderson BJ, Tyndel S, Bankhead C, Watson E, Clements A, *et al.* Predictors of breast cancer-related distress following mammography screening in younger women on a family history breast screening programme. *Psychooncology* 2008;**17**:1180–8. http://dx.doi.org/10.1002/pon.1355

103. Brett J, Austoker J, Ong G. Do women who undergo further investigation for breast screening suffer adverse psychological consequences? A multi-centre follow-up study comparing different breast screening result groups five months after their last breast screening appointment. *J Public Health* 1998;**20**:396–403. http://dx.doi.org/10.1093/oxfordjournals.pubmed.a024793

104. Ong G, Austoker J, Brett J. Breast screening: adverse psychological consequences one month after placing women on early recall because of a diagnostic uncertainty. A multicentre study. *J Med Screen* 1997;**4**:158–68.

105. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. *Br J Cancer* 1989;**60**:781–4. http://dx.doi.org/10.1038/bjc.1989.359

106. Bull AR, Campbell MJ. Assessment of the psychological impact of a breast screening programme. *Br J Radiol* 1991;**64**:510–15. http://dx.doi.org/10.1259/0007-1285-64-762-510

107. Clements A, Tyndel S, Henderson B, Brain K, Watson E, Austoker J, *et al.* 'More positive about mammography' – reactions of women to a false positive recall: a qualitative study of women at risk of familial breast cancer. *Breast Cancer Res* 2008;**10**:90. http://dx.doi.org/10.1186/bcr1974

108. O'Sullivan I, Sutton S, Dixon S, Perry N. False positive results do not have a negative effect on reattendance for subsequent breast screening. *J Med Screen* 2001;**8**:145–8. http://dx.doi.org/10.1136/jms.8.3.145

109. Orton M, Fitzpatrick R, Fuller A, Mant D, Mlynek C, Thorogood M. Factors affecting women's response to an invitation to attend for a second breast cancer screening examination. *Br J Gen Pract* 1991;**41**:320–2.

110. Ong G, Austoker J. Recalling women for further investigation of breast screening: women's experiences at the clinic and afterwards. *J Public Health Med* 1997;**19**:29–36. http://dx.doi.org/10.1093/oxfordjournals.pubmed.a024582

111. Ong G, Austoker J, Brouwer A. Evaluation of the written information sent to women who are called back for further investigation of breast screening in the UK. *Health Education Journal* 1996;**55**:413–29. http://dx.doi.org/10.1177/001789699605500407

112. Austoker J, Ong G. Written information needs of women who are recalled for further investigation of breast screening: results of a multicentre study. *J Med Screen* 1994;**1**:238–44.

113. Smith S, Botha JL, Goosey R, Daintith H. Audit of user satisfaction with the Leicestershire Breast Screening Service – women attending for assessment of abnormal mammograms. *J Public Health Med* 1991;**13**:166–71.

114. Ong G, Austoker J, Brouwer A, Programme NHSB. *Guidelines on improving the quality of the written information sent to women who are recalled for assessment. NHSBSP Publication no. 38*. NHSBSP; 1998.

115. Meldrum P, Turnbull D, Dobson HM, Colquhoun C, Gilmour WH, McIlwaine GM. Tailored written invitations for second round breast cancer screening: a randomised controlled trial. *J Med Screen* 1994;**1**:245–8.

116. Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. *J Med Screen* 2004;**11**:39–44. http://dx.doi.org/10.1258/096914104772950745

117. Brodersen J, Thorsen H, Kreiner S. Validation of a condition-specific measure for women having an abnormal screening mammography. *Value Health* 2007;**10**:294–304. http://dx.doi.org/10.1111/j.1524-4733.2007.00184.x

118. Brodersen J, McKenna SP, Doward LC, Thorsen H. Measuring the psychosocial consequences of screening. *Health Qual Life Outcomes* 2007;**5**:3. http://dx.doi.org/10.1186/1477-7525-5-3

119. Brodersen J, Thorsen H, Anderson JS. Short-term and long-term psychosocial consequences of false positive screening mammography – development of two new questionnaire based on the psychological consequences questionnaire (the PCQ). *Value Health* 2004;**7**:786–7. http://dx.doi.org/10.1016/S1098-3015(10)66105-0

120. Wardle J, Pope R. The psychological costs of screening for cancer. *J Psychosom Res* 1992;**36**:609–24. http://dx.doi.org/10.1016/0022-3999(92)90051-3

121. Chalmers K, Thomson K. Coming to terms with the risk of breast cancer: perceptions of women with primary relatives with breast cancer. 1996;**6**:256–82.

122. Scaf-Klomp W, van Sonderen FL, Stewart R, van Dijck JA, van den Heuvel WJ. Compliance after 17 years of breast cancer screening. *J Med Screen* 1995;**2**:195–9.

123. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, *et al.* Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;**7**(27).

124. Mays N, Pope C. Qualitative research: rigour and qualitative research. *BMJ* 1995;**311**:109–12. http://dx.doi.org/10.1136/bmj.311.6997.109

125. Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. *Qual Health Res* 1998;**8**:341–51. http://dx.doi.org/10.1177/104973239800800305

126. Noblit GW, Hare RD. *Meta-ethnography: synthesizing qualitative studies*. Newbury Park, CA: Sage; 1988.

127. Britten N, Campbell R, Pope C, Donovan J, Morgan M, Pill R. Using meta ethnography to synthesise qualitative research: a worked example. *J Health Serv Res Policy* 2002;**7**:209–15. http://dx.doi.org/10.1258/135581902320432732

128. Garside R, Britten N, Stein K. The experience of heavy menstrual bleeding: a systematic review and meta-ethnography of qualitative studies. *J Adv Nurs* 2008;**63**:550–62. http://dx.doi.org/10.1111/j.1365-2648.2008.04750.x

129. Rodgers M, Sowden A, Pettigrew M, Arai L, Roberts H, Britten N, *et al. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews*. 2009. URL: http://evi.sagepub.com (cited 12 June 2009).

130. Oliver S, Harden A, Rees R, Shepherd J, Brunton G, Oakley A. Young people and mental health: novel methods for systematic review of research on barriers and facilitators. *Health Educ Res* 2008;**23**:770–90. http://dx.doi.org/10.1093/her/cym038

Appendix 1 Protocol

TECHNOLOGY ASSESSMENT REPORT COMMISSIONED BY THE NETSCC HTA PROGRAMME

HTA 09/145

Plain English Summary

In the UK women aged 50 to 70 years old are invited to come for mammography screening every 3 years. About 5% of these are recalled for further investigation. After follow-up it is found that about 82% of recalled women had nothing wrong with them (false-positives). However, the experience of being unnecessarily recalled can be distressing, not just in the short-term but may lead to enduring anxiety and affect attendance at future routine mammography screening. The purpose of this systematic review is to find out what the research evidence is for medium and long-term effects of having a false-positive mammogram on mental health and behaviour, whether some groups of women are more likely to be adversely affected than others and if there are ways of reducing the negative effects of being recalled when you are in fact well.

Decision problem

The purpose of this technology assessment is to conduct a systematic review, to identify the psychological and behavioural consequences following false-positive screening mammogram results that affect women and any evidence for the effectiveness of interventions designed to reduce these. In particular we will be looking at whether the psychological and behavioural consequences or the effectiveness of specific interventions differ in different groups of women.

This research is necessary because of the large number of false-positive results that come from routine mammography screening. In the UK women aged 50-70 years, on population registers, are invited for mammography every 3 years through the NHS Breast Screening Programme (NHSBSP). Around two million women were screened by the NHSBSP in 2007/8 and of these 95,006 (5%) were recalled for further investigation; 16,735 cancers were detected leaving 78,271(82%) false-positive recalls.[2]

Quantitative observational studies looking at the psychological and behavioural consequences of false-positive mammograms show conflicting results. Some studies indicate that, while women show increased distress between receiving the information about the need for a follow-up appointment and receiving the all-clear, in the longer term their anxieties about breast cancer and mammography are not increased.[53–55] Other studies report that there are long-term adverse psychological consequences to receiving a false-positive mammogram.[56,57,59,103] The outcomes of studies looking at whether having false-positive results affects future attendance at breast screening appointments is similarly conflicted.[58–61]

A quantitative systematic review in 2007 by Brewer and colleagues found that the impact of a false-positive mammogram on subsequent screening attendance varied with nationality; although, the reasons for this were unclear. They also reported a varying impact on long-term psychological distress, anxiety and depression, and other behaviours such as frequency of breast self-examination.[62] However, their review did not report the reasons for this variation in response. Furthermore, Brewer and colleague's review found no statistically sound studies that investigated whether anxiety over a false-positive mammogram directly affects whether women return for routine screening or increase breast self-examination. There was little evidence about the effects on quality of life or trust of healthcare services and no evidence about whether

women who felt anxious after a false-positive screening result replaced routine screening attendance with breast self-examination.[62] We also do not know what meanings women attribute to a false-positive mammogram or how these may determine their behaviour when invited for further routine mammogram screening as qualitative evidence is lacking.

Therefore, there is uncertainty about the psychological impact of false-positive mammograms on women. We do not know what the mediators are of negative psychological and behavioural outcomes which may affect attendance at future mammography screening. There is a need to answer these questions to identify and evaluate studies of interventions to treat the effect of false-positive results, and identify whether these effects differ in women from different backgrounds. The answers will have important policy implications for the NHS in the provision of breast cancer screening services.

The questions that this systematic review will answer are:

1. What evidence is there for medium or long-term adverse psychological consequences of false-positive screening mammograms?
   (a) Do the types of psychological consequences differ between different groups of women?
2. Are there interventions that reduce adverse psychological consequences?

For question one the population will be women who have received a false-positive result from routine mammogram screening in the UK and invited for further assessment. Where studies include a comparator this will be women who had a routine screening mammogram but who had a normal mammogram and were not invited for further assessment. A range of outcomes, including qualitative, will be considered that report psychological and behavioural measures over the medium and long-term. Where data permit, subgroup analyses will be conducted of different groups of women (including socio-economic status and ethnic group).

For question two the population and the outcomes will be the same as question one. The interventions will be those delivered to individuals to address the adverse psychological consequences of a false-positive mammogram result, including attendance at future routine breast screening. Where there are comparators this will be an absence of an individualized intervention in the same population. Where data permit, subgroup analyses will be conducted of different groups of women (including socio-economic status and ethnic group).

It is intended that this should be a wide systematic review considering a range of study types including uncontrolled studies and qualitative research but excluding individual case studies. Recommendations will be made for future primary research.

Methods for selection of evidence of clinical effectiveness

A systematic review will be conducted using the principles of the NHS Centre for Reviews and Dissemination[77] including those for non-randomized and qualitative studies.[123]

Inclusion criteria

Question	Criteria	Specification	Notes
1 and 2	Population	Women who have received a positive result from routine mammogram screening in the UK and have been invited for further assessment which shows that they do not have breast cancer	Where data permit we will look at sub groups including socio-economic status, and ethnic group
2	Intervention	Those interventions delivered to individuals to address the adverse psychological and behavioural consequences of a false-positive mammogram result.	These are individual interventions not group ones
1	Comparator	Women who have received a negative (normal) result from routine mammogram screening in the UK.	
2	Comparator	An absence of an individual intervention in the same population	
1 and 2	Outcomes	Psychological and behavioural outcomes and those from qualitative studies	Including subsequent attendance at routine mammography screening and quality of life
1 and 2	Setting	UK	Secondary care
1 and 2	Study design	Systematic reviews, randomized, non-randomized, observational and qualitative studies	We will not consider individual case studies
1 and 2	Length of follow-up	At least one month from the 'all-clear'	Measured over the medium to long-term, i.e. not the immediate response to receiving a false-positive result
1 and 2	Language	English language only	Non English language papers will be included in the searches and screened, so that the number of potentially includable foreign language papers is known

Exclusion criteria

The following types of studies will be excluded: narrative reviews, editorials, opinion pieces, non-English language papers, individual case studies, and studies only reported as posters or by abstract where there is insufficient information to assess the quality of the study.

Search strategy

Refer to Appendix 1 for the draft search strategy for MEDLINE.

The search strategy will comprise the following main elements:

- Searching of electronic bibliographic databases.
- Internet searches.
- Scrutiny of references of included studies.
- Contacting experts in the field.

Databases will include:

MEDLINE, EMBASE, The Cochrane Library, PsychLIT, CINAHL EBSCO, Web of Science, Science Citation Index Expanded, Conference Proceedings Citation Index, Sociological Abstracts, Applied Social Sciences Index, Sociological Abstracts, Applied Social Sciences Index and International Bibliography of the Social Sciences.

Study selection

Based on the above inclusion/exclusion criteria, papers will be selected for review from the titles and abstracts generated by the search strategy. This will be done independently by two reviewers; discrepancies will be resolved by discussion, with the involvement of a third reviewer if necessary. Although non-English language papers will not be included in the systematic review due to resource limitations, they will be identified and any that meet the other inclusion criteria will be recorded with their language noted as the reason for their exclusion. Retrieved papers will again be reviewed and selected against the inclusion criteria by the same independent process.

Data extraction

Data will be extracted from included studies by one reviewer using a standardised data extraction form and checked by another reviewer. Authors of studies will be contacted to provide missing information, as necessary.

Quality assessment

Quantitative studies will be assessed for internal and external validity according to criteria suggested by the updated NHS CRD Report No.4, according to study type.[77,78] Qualitative studies will have their quality assessed using a standard assessment tool, e.g. Mays and Pope 1995[124] and Popay and colleagues 1998,[125] a number of these will be piloted to assess their suitability for the task.

Methods for analysis and synthesis of evidence of clinical effectiveness

Quantitative analysis and synthesis

Studies were assessed for internal and external validity according to criteria suggested by the updated NHS CRD Report No.4, according to study type.[77,78] The quality of systematic reviews was evaluated using the PRISMA statement.[79] Individual RCTs were appraised with the CONSORT statement[80] and individual observational studies with STROBE guidelines.[81]

Qualitative analysis and synthesis

These studies will be analysed using meta-ethnography[126–128] supported by Atlas.ti6 software. Here the included studies' results are translated into one another, while preserving their original meaning, with an inductive and interpretive approach to allow comparison between them. Authors' interpretation of the primary study findings become the data, which are translated across studies by the reviewers to produce a synthesis of meaning allowing the production of higher order concepts.

Combined synthesis of quantitative and qualitative evidence

The results of the quantitative and qualitative analyses will undergo narrative synthesis to construct an explanatory framework.[129,130] In this method both types of data analysis undergo a further narrative synthesis of their combined data through a process of developing an explanatory theory, undertaking a preliminary synthesis, looking at the relationships between and within studies and evaluating the robustness of the synthesis.

Expertise in this TAR team

People

Name	Institution	Expertise
Mrs Mary Bond	PenTAG, University of Exeter	Systematic reviewing, psychology and project management
Dr Toby Pavey	PenTAG, University of Exeter	Systematic reviewing
Mrs Karen Welch	Karen Welch Information Consultancy	Information Specialist
Mr Chris Cooper	PenTAG, University of Exeter	Information Specialist
Dr Ruth Garside	PenTAG, University of Exeter	Qualitative evidence synthesis
Prof. Chris Hyde	PenTAG, University of Exeter	Diagnostics and public health

In addition to the research team, we will be receiving expert clinical advice from Dr Russell Davies Consultant Breast Radiologist (Royal Devon and Exeter Foundation Trust), Gillian Gray (Breast Care nurse Royal Devon and Exeter Foundation Trust), Dr Jim Steel Consultant Breast Radiologist and Prof Carl Roobottom, Consultant Radiologist (both at Derriford Hospital, Plymouth), Jenny Hewison Professor of the Psychology of Healthcare, from the University of Leeds. We have two patient representatives, Kate Blackmore and Sue Milward who have both had experience of having a false-positive mammogram to advise us on the patient perspective.

TAR centre – PenTAG

This project is being conducted by The Peninsula Technology Assessment Group (PenTAG), which is part of the Institute of Health Service Research at the Peninsula Medical School, University of Exeter. PenTAG was established in 2000 and carries out independent Health Technology Assessments for the UK HTA Programme and other local and national decision-makers including NICE. The group is multi-disciplinary and draws on individuals' backgrounds in public health, health services research, computing and decision analysis, systematic reviewing, psychology, statistics and health economics. The Institute of Health Service Research is made up of discrete but methodologically related research groups, among which Health Technology Assessment is a strong and recurring theme.

Contributions of team members

Name	Job title	Contribution
Mary Bond	Research Fellow in Health Technology Assessment	Providing project management. Writing the protocol. Conducting the systematic review. Writing and editing the report.
Toby Pavey	Research Fellow in Health Technology Assessment	Second reviewing the titles, abstracts and papers for the systematic review.
Karen Welch	Information Specialist	Writing and running the search strategies for the systematic review
Chris Cooper	Information Specialist	Writing and running the search strategies for the systematic review
Ruth Garside	Senior Research Fellow	Overseeing qualitative evidence synthesis
Chris Hyde	Professor of Public Health and Clinical Epidemiology	Director of the project and guarantor of the report. Contributing to editing the report.

Competing interests of authors

None.

Timetable and project milestones

Milestones	November	December	January	February	March	April	May	June	July	August	September	October
Conduct searches												
Select studies												
Analysis												
Writing up												

Appendix 2 Search strategy

All searches combined results database summary.

Database	Search date	Hits
MEDLINE Ovid Scoping Search: 1950–present	8 October 2010 and 25 January 2011	2357
MEDLINE In-Process & Other Non-Indexed Citations Ovid		98
EMBASE Classic + EMBASE Ovid: 1947–2010	22 November 2010 and 25 January 2011	2672
HMIC Ovid		58
Source: U.K. Department of Health, Nuffield Institute for Health (Leeds University Library), King's Fund Library; 1983–2010	21 November 2010 and 25 January 2011	2
CCRCT: all years	22 November 2010 and 25 January 2011	111
Cochrane CDSR: all years	22 November 2010 and 25 January 2011	7
CRD DARE: all years	22 November 2010 and 25 January 2011	17
CRD HTA: all years	22 November 2010	8
Cochrane Methodology: all years	22 November 2010, 9 December 2010 and 25 January 2011	4
Web of Science: all years	3 December 2010 and 25 January 2011	406
Science Citation Index expanded: 1970–present		
SSCI: 1970–present		
CPCI-S:1990–present		
CPCI-SSH:1990–present		
PsychINFO EBSCO: 1887–2010	6 December 2010 and 25 January 2011	152
CINAHL EBSCO: 1981–2010	6 December 2010 and 25 January 2011	260
Sociological Abstracts CSA Illumina: 1952–2010	6 December 2010 and 25 January 2011	13
IBSS CSA Illumina: 1951–2010	6 December 2010 and 25 January 2011	22
BNI	25 January 2011	13
SPP	25 January 2011	3
ASSIA	25 January 2011	78
CRD	25 January 2011	245
Total		6526

The MEDLINE search strategy was translated and run in:

Database	Search date	Hits
MEDLINE Ovid Scoping Search: 1950–present	8 October 2010	559
MEDLINE In-Process & Other Non-Indexed Citations Ovid		17
EMBASE Classic + EMBASE Ovid: 1947–2010	22 November 2010	500
HMIC Ovid		2
Source: U.K. Department of Health, Nuffield Institute for Health (Leeds University Library), King's Fund Library; 1983–2010	21 November 2010	2
CCRCT: all years	22 November 2010	68
Cochrane CDSR: all years	22 November 2010	7
CRD DARE: all years	22 November 2010	17
CRD HTA: all years	22 November 2010	8
Cochrane Methodology: all years	22 November 2010 and 9 December 2010	4
Web of Science: all years	3 December 2010	129
Science Citation Index expanded: 1970–present		
SSCI: 1970–present		
CPCI-S:1990–present		
CPCI-SSH: 1990–present		
PsychINFO EBSCO: 1887–2010	6 December 2010	28
CINAHL EBSCO: 1981–2010	6 December 2010	99
Sociological Abstracts CSA Illumina: 1952–2010	6 December 2010	1
IBSS CSA Illumina: 1951–2010	6 December 2010	0

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
MEDLINE Ovid Scoping Search 1950–current Searched on 8 October 2010	1. exp mammography/ae, px 2. exp mammography/ 3. FFDM.tw. 4. (mammogram* or mammograph*).tw. 5. (breast adj2 screen*).tw. 6. (breast adj2 scan*).tw. 7. "National Health Service Breast Screening Programme".tw. 8. NHSBSP.tw. 9. UK breast screen* program*.tw. 10. NHS breast screen* program*.tw. 11. Mass Screening/ 12. exp Breast Neoplasms/ 13. 11 and 12 14. 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 13 15. False Positive Reactions/ 16. (false* adj3 positive*).tw. 17. "false-positive".tw. 18. "false-positives".tw. 19. (false adj3 test*).tw. 20. (false adj3 retest*).tw. 21. (retest* adj3 negative).tw. 22. diagnostic uncertaint*.tw. 23. or/15-22 24. exp Stress, Psychological/ 25. exp anxiety/ 26. exp fear/ 27. exp Depression/ 28. exp Emotions/ 29. Psychophysiologic Disorders/ 30. exp Psychology/ 31. exp Health Behavior/ 32. exp Behavior/ 33. exp attitude/ 34. Motivation/ 35. Decision Making/ 36. exp "Quality of Life"/ 37. Health Knowledge, Attitudes, Practice/ or Attitude to Health/ or Patient Satisfaction/ or Patient Participation/ or Consumer Participation/ or Consumer Satisfaction/ or Sick Role/ or "Patient Acceptance of Health Care"/ 38. exp Affect/ 39. exp Affective Symptoms/ 40. (accept* or adhere* or affect* or anger* or anxiety or anxious or alarm* or attitude* or appetite or behavior* or behaviour* or belief* or believe* or comply or complian* or concordance or coping or concern* or confusion or confused or consequence* or consequential or conflict or cultural*).tw.	559

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	41. (demotivated or demotivation* or de-motivated or de-motivation* or disconcert* or depression or depressed or distress* or deleterious or disappointment or emotion* or ethnic* or ethnol* or experienc* or fear* or fright* or harm* or mental* or mistrust* or mood* or motivated or motivation* or nightmare* or perception* or perceive* or psychological or psychologically or psychology or psychosocial or reattend* or social*).tw. 42. "quality of life".tw. 43. (relief or relieved or risk*).tw. 44. (sleep or stress* or terror or terrified or trust* or mistrust*).tw. 45. (worry or worried).tw. 46. (wellbeing or "well-being" or "well being").tw. 47. or/24-46 48. exp Intervention Studies/ 49. exp Questionnaires/ 50. psychological tests/ or psychometrics/ or models psychological/ 51. Patient Education as Topic/ 52. health education/ or health promotion/ or health knowledge/ 53. decision aid/ or decision support techniques/ 54. Educational Technology/ 55. audiovisual aids/ 56. telehealth/ or telemedicine/ or telecommunication/ 57. social support/ or self help groups/ or support groups/ 58. exp communication/ 59. persuasive communication/ 60. exp counselling/ 61. interviews as topic/ 62. evaluation studies as topic/ 63. qualitative research/ or program evaluation/ or process evaluation/ 64. focus groups/ 65. nursing methodology research/ 66. intervention*.tw. 67. (qualitative* or findings or evaluat* or synthes?s or meta-synthesis* or meta synthesis* or metasynthesis or meta-ethnograph* or metaethnograph* or meta ethnograph* or meta-study or metastudy or meta study or systematic* or "technology assessment" or sampl* or study or studies or observation* or research or discourse* or analys?s or humanistic or biographical or biography or narrative*).tw. 68. (support* or literature or booklet* or leaflet* or pamphlet* or letter* or video* or podcast* or telephon* or transtelephon*).tw. 69. (questionnaire* or interview* or discuss* or feedback or personalised or personalized or assessment* or reassurance or reassur*).tw. 70. (counsel* or education* or "informed choice" or "informed choices").tw. 71. "in person".tw. 72. (peer* adj5 (support* or group*)).tw. 73. ("expert patients" or "expert patients").tw. 74. (social adj network*).tw. 75. "emotional support".tw.	

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	76. "family support".tw. 77. focus group*.tw. 78. ("one to one" or "one on one").tw. 79. ((patient* or consumer* or recipient* or client* or individual*) adj5 (communicat* or counsel* or inform* or education* or choice or discuss* or decision* or decide* or participat* or preference* or navigat*)).tw. 80. ((patient* or consumer*or recipient* or client* or individual*) adj5 (tailor* or personal*)).tw. 81. ((personal or interpersonal* or individual*) adj5 (decision* or choice* or preference* or participat* or preference*)).tw. 82. ((tailor* or individual* or personal*) adj5 message*).tw. 83. ((allocat* or allot* or assign* or divid*) adj5 (condition* or experiment* or intervention* or treatment* or therap* or control* or group*)).tw. 84. or/48-83 85. 1 and 23 and 84 86. 14 and 23 and 47 and 84 87. 85 or 86 88. 1 and 23 89. 14 and 23 and 47 90. 88 or 89 91. limit 90 to ("qualitative studies (sensitivity)" or "qualitative studies (specificity)" or "qualitative studies (optimized)") 92. limit 90 to systematic reviews 93. limit 90 to (case reports or clinical trial, all or clinical trial or comparative study or controlled clinical trial or evaluation studies or government publications or guideline or meta analysis or multicenter study or patient education handout or practice guideline or randomized controlled trial or "review" or "scientific integrity review" or technical report or twin study or validation studies) 94. 87 or 91 or 92 or 93 95. 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 64 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 96. 48 or 49 or 61 or 62 or 63 or 65 or 66 or 67 97. 14 and 23 and 47 and 95 and 96 98. 1 and 23 and 96 99. 14 and 23 and 47 and 96 100. 94 or 97 or 98 or 99 101. 94 or 100	
MEDLINE In-Process & other Non-Indexed Citations Ovid	Adapted from MEDLINE Search	17

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
EMBASE Classic + EMBASE Ovid 1947–2010 Searched 22 November 2010	Database: EMBASE Classic + EMBASE <1947 to 2010 Week 46> Search strategy: 1. exp MAMMOGRAPHY/ (33,626) 2. FFDM.tw. (140) 3. (mammogram* or mammograph*).tw. (24,341) 4. (breast adj2 screen*).tw. (6407) 5. (breast adj2 scan*).tw. (245) 6. (breast adj2 detect*).tw. (5038) 7. "National Health Service Breast Screening Programme".tw. (96) 8. NHSBSP.tw. (76) 9. UK breast screen* program*.tw. (42) 10. NHS breast screen* program*.tw. (86) 11. mass screening/ (43,653) 12. exp breast cancer/ (208,636) 13. 11 and 12 (873) 14. or/1-10,13 (41,820) 15. false positive result/ (6554) 16. (false* adj3 positive*).tw. (44,807) 17. "false-positive".tw. (36,608) 18. "false-positives".tw. (8293) 19. "falsely-positive".tw. (855) 20. (false adj3 test*).tw. (3821) 21. (false adj3 retest*).tw. (12) 22. (retest* adj3 negative).tw. (115) 23. diagnostic uncertaint*.tw. (743) 24. or/15-23 (48,970) 25. 14 and 24 (1368) 26. exp stress/ (158,295) 27. PSYCHOLOGICAL ASPECT/ or PSYCHOLOGICAL WELL BEING/ (413,535) 28. PSYCHOSOCIAL ADJUSTMENT TO ILLNESS SCALE/ or exp "PSYCHOLOGICAL AND PSYCHOSOCIAL PHENOMENA"/ (1,951,140) 29. exp ANXIETY/ or ANTICIPATORY ANXIETY/ (89,596) 30. exp FEAR/ (114,051) 31. exp DEPRESSION/ or REACTIVE DEPRESSION/ (244,313) 32. exp EMOTION/ (282,458) 33. EMOTIONAL STRESS/ (9740) 34. mental stress/ (46,546) 35. exp behavior/ (1,772,614) 36. exp ATTITUDE/ (384,744) 37. exp MOTIVATION/ (54,973) 38. PATIENT PARTICIPATION/ (12,113) 39. decision making/ (106,459) 40. exp "quality of life"/ (165,080) 41. exp patient satisfaction/ (61,974)	500

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	42. (accept* or adhere* or affect* or anger* or anxiety or anxious or alarm* or attitude* or appetite or behavior* or behaviour* or belief* or believe* or comply or complian* or concordance or cope* or coping or concern* or confusion or confused or consequence* or consequential or conflict or cultural*).tw. (3,217,709)	
	43. (demotivated or demotivation* or de-motivated or de-motivation* or disconcert* or depression or depressed or distress* or deleterious or disappointment or emotion* or ethnic* or ethnol* or experienc* or fear* or fright* or harm* or mental* or mistrust* or mood* or motivated or motivation* or nightmare* or perception* or perceicve* or psychological or psychologically or psychology or psychosocial or reattend* or social*).tw. (1,974,157)	
	44. "quality of life".tw. (128,403)	
	45. (relief or relieved or risk*).tw. (1,252,141)	
	46. (sleep or stress* or terror or terrified or trust* or mistrust*).tw. (611,497)	
	47. (worry or worried or worries or wellbeing or "well-being" or "well being").tw. (48,645)	
	48. or/26-47 (6,645,255)	
	49. 25 and 48 (635)	
	50. qualitative stud$.mp. (11,374)	
	51. nursing methodology research.mp. (11,841)	
	52. questionnaire/ (262,746)	
	53. focus group$.mp. (14,900)	
	54. discourse analysis.mp. (846)	
	55. content analysis.mp. (8358)	
	56. ethnographic research.mp. (570)	
	57. ethnological research.mp. (6)	
	58. ethnonursing research.mp. (27)	
	59. constant comparative method.mp. (597)	
	60. qualitative validity.mp. (83)	
	61. purposive sample.mp. (1167)	
	62. observational method$.mp. (959)	
	63. field stud$.mp. (10,752)	
	64. theoretical sampl$.mp. (293)	
	65. phenomenology/ (5404)	
	66. phenomenological research.mp. (218)	
	67. life experience$.mp. (2658)	
	68. cluster sampl$.mp. (2894)	
	69. ethnonursing.af. (78)	
	70. ethnograph$.mp. (4843)	
	71. phenomenol$.af. (15,839)	
	72. grounded theory.mp. (4137)	
	73. (grounded adj (theor$ or study or studies or research or analys?s)).af. (4231)	
	74. (life stor$ or women$ stor$).tw. (658)	
	75. (emic or etic or hermeneutic$ or heuristic$ or semiotic$).af. or (data adj1 saturat$).tw. or participant observ$.tw. (11,156)	
	76. (action research or cooperative inquir$ or co operative inquir$ or co-operative inquir$).mp. (1979)	

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	77. (humanistic or existential or experiential or paradigm$).mp. (78,597)	
	78. (field adj (study or studies or research)).tw. (11,218)	
	79. human science.tw. (215)	
	80. biographical method.tw. (14)	
	81. purposive sampl$.af. (1846)	
	82. theoretical sampl$.af. (293)	
	83. ((purpos$ adj4 sampl$) or (focus adj group$)).af. (19,032)	
	84. (account or accounts or unstructured or "open-ended" or "open ended" or text$ or narrative$).mp. (395,335)	
	85. (life world or life-world or conversation analys?s or personal experience$ or theoretical saturation).mp. (22,256)	
	86. lived experience$.tw. (1610)	
	87. life experience$.tw. (2647)	
	88. cluster sampl$.mp. (2894)	
	89. questionnaire$.tw. (266,283)	
	90. content analysis.af. (8419)	
	91. thematic analysis.af. (2239)	
	92. constant comparative.af. (1043)	
	93. ((discourse$ or discuss$ or discurs$) adj3 analys?s).tw. (6735)	
	94. (constant adj (comparative or comparison)).af. (1486)	
	95. narrative analys?s.af. (358)	
	96. narrative synthes?s.af. (152)	
	97. observational study/ (16,809)	
	98. intervention study/ (9965)	
	99. intervention*.tw. (471,470)	
	100. evaluation/ (159,571)	
	101. evaluat*.tw. (2,055,889)	
	102. psychometry/ (31,871)	
	103. PSYCHOLOGICAL RATING SCALE/ or PSYCHOLOGICAL MODEL/ (27,953)	
	104. PATIENT EDUCATION/ or HEALTH EDUCATION/ (137,523)	
	105. educational technology/ or health promotion/ (55,616)	
	106. telecommunication/ or TELEHEALTH/ or telemedicine/ (19,465)	
	107. SOCIAL SUPPORT/ or SUPPORT GROUP/ (46,011)	
	108. self help/ (9686)	
	109. communication.mp. (293,687)	
	110. exp COUNSELLING/ (82,605)	
	111. qualitative research/ (11,352)	
	112. program evaluation.mp. or health care quality/ (143,551)	
	113. (support* or literature or booklet* or leaflet* or pamphlet* or letter* or video* or podcast* or telephon* or transtelephon*).tw. (1,595,815)	
	114. (questionnaire* or interview* or discuss* or feedback or personalised or personalized or assessment* or reassurance or reassur*).tw. (2,123,205)	
	115. (counsel* or education* or "informed choice" or "informed choices").tw. (378,757)	
	116. "in person".tw. (3324)	

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	117. (peer* adj5 (support* or group*)).tw. (5061) 118. ("expert patients" or "expert patients").tw. (70) 119. (social adj network*).tw. (5102) 120. "emotional support".tw. (3119) 121. "family support".tw. (2620) 122. ((patient* or consumer* or recipient* or client* or individual*) adj5 (communicat* or counsel* or inform* or education* or choice or discuss* or decision* or decide* or participat* or preference* or navigat*)).tw. (227961) 123. ((patient* or consumer*or recipient* or client* or individual*) adj5 (tailor* or personal*)).tw. (28,228) 124. ((tailor* or individual* or personal*) adj5 message*).tw. (629) 125. ((allocat* or allot* or assign* or divid*) adj5 (condition* or experiment* or intervention* or treatment* or therap* or control* or group*)).tw. (224,681) 126. or/50-125 (6,376,617) 127. 49 and 126 (426) 128. limit 49 to (clinical trial or randomized controlled trial or controlled clinical trial or multicenter study) (78) 129. limit 49 to (evidence based medicine or meta analysis or outcomes research or "systematic review") (37) 130. limit 49 to ("qualitative studies (1 term high sensitivity)" or "qualitative studies (1 term high specificity)" or "qualitative studies (1 term min difference)" or "qualitative studies (2 or more terms high sensitivity)" or "qualitative studies (2 or more terms high specificity)" or "qualitative studies (2 or more terms min difference)") (211) 131. 127 or 128 or 129 or 130 (498) 132. (psychological adj5 (consequence* or harm* or sequel*)).tw. (3753) 133. 25 and 132 (17) 134. 131 or 133 (500)	
HMIC Ovid (Health Management Information Consortium) Source: U.K. Department of Health, Nuffield Institute for Health (Leeds University Library), King's Fund Library 1983–2010 Searched 21 November 2010	Free text from MEDLINE search	*n* = 2 added (23 results downloaded would not import so cross-checked records manually, only 2 unique so only these added manually to the database)

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
CCRCT	Save date: 22 November 2010	68
All Years	ID Search	
Searched 22 November 2010	#1 MeSH descriptor False Positive Reactions explode all trees #2 false* positive* #3 (#1 OR #2) #4 MeSH descriptor Mammography, this term only #5 (mammogram* or mammograph* or breast screen* or breast scan*) #6 (National Health Service Breast Screen* Programme* or NHSBSP or FFDM) #7 breast cancer screen* or breast cancer scan* #8 (breast neoplasm* scan* or breast neoplasm* screen*) #9 (#4 OR #5 OR #6 OR #7 OR #8) #10 (#3 AND #9) #11 (accept* or adhere* or affect* or anger* or anxiety or anxious or alarm* or attitude* or appetite or behavior* or behaviour* or belief* or believe* or comply or complian* or concordance or coping or concern* or confusion or confused or consequence* or consequential or conflict or cultural*) #12 (demotivated or demotivation* or de-motivated or de-motivation* or disconcert* or depression or depressed or distress* or deleterious or disappointment or emotion* or ethnic* or ethnol* or experienc* or fear* or fright* or harm* or mental* or mistrust* or mood* or motivated or motivation* or nightmare* or perception* or perceive* or psychological or psychologically or psychology or psychosocial or reattend* or social*) #13 quality of life #14 (sleep or stress* or terror or terrified or trust* or mistrust* or worry or worries or worried) #15 (wellbeing or well-being or well being) #16 (#11 OR #12 OR #13 OR #14 OR #15) #17 (#10 AND #16) #18 MeSH descriptor Mammography explode all trees with qualifier: PX #19 (#17 OR #18) #20 psychological near false #21 (#9 AND #20) #22 (#19 OR #21)	
Cochrane CDSR All years Searched 22 November 2010	As above	7 selected (out of 25 records)

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
CRD DARE All years Searched 22 November 2010 CRD HTA All years Searched 22 November 2010	#1 mammogram* OR mammograph* 353 #2 breast NEAR screen* 321 #3 breast NEAR scan* 60 #4 MeSH Mammography EXPLODE 1 315 #5 NHSBSP OR FFDM 5 #6 #1 or #2 or #3 or #4 or #5 583 #7 MeSH False Positive Reactions EXPLODE 1 119 #8 MeSH False Positive Reactions EXPLODE 1 119 #9 false* NEAR positive* 736 #10 "false-positive" 556 #11 "false-positives" 220 #12 false NEAR test* 398 #13 false NEAR retest* 2 #14 retest* NEAR negative 5 #15 negative NEAR retest* 5 #16 diagnos* NEAR uncertaint* 181 #17 #7 or #9 or #10 or #11 or #12 or #13 or #14 or #16 995 #18 #6 and #17 87 #19 (accept* OR adhere* OR affect* OR anger* OR anxiety OR anxious OR alarm* OR attitude* OR appetite OR behavior* OR behaviour* OR belief* OR believe* OR comply OR complian* OR concordance OR coping OR concern* OR confusion OR confused OR consequence* OR consequential OR conflict OR cultural*) 14608 #20 (demotivated OR demotivation* OR de-motivated OR de-motivation* OR disconcert* OR depression OR depressed OR distress* OR deleterious OR disappointment OR emotion* OR ethnic* OR ethnol* OR experienc* OR fear* OR fright* OR harm* OR mental* OR mistrust* OR mood* OR motivated OR motivation* OR nightmare* OR perception* OR perceive* OR psychological OR psychologically OR psychology OR psychosocial OR reattend* OR social*) 12173 #21 "quality of life" 5283 #22 (sleep OR stress* OR terror OR terrified OR trust* OR mistrust*) 2165 #23 (worry OR worried) 50 #24 (wellbeing OR "well-being" OR "well being") 568 #25 MeSH Mammography EXPLODE 1 315 #26 #19 or #20 or #21 or #22 or #23 or #24 22424 #27 #18 and #26 63 #28 MeSH Mammography QUALIFIERS PX EXPLODE 1 4 #29 #27 or #28 66	17 DARE; 8 HTA
Cochrane Methodology All years searched 22 November 2010 and 9 December 2010	4 references manually added to Reference Manager	4

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)		Number of results
Web of Science Searched 3 December 2010 Cross searched: databases Timespan = all available years SCI-Expanded: 1970–present SSCI:1970–present CPCI-S: 1990–present CPCI-SSH: 1990–present	#1	25,094 TS=(mammogram* or mammograph*)	*n* = 129 (123 from line 36 + extra 6 chosen out of line 37)
	#2	84 TS=(NHSBP or "National Health Service Breast Screening Programme")	
	#3	243 TS=(UK breast screen* program*)	
	#4	115 TS=(NHS breast screen* program*)	
	#5	34,318 TS=(false* positive*)	
	#6	12,091 TS=(diagnos* uncertain*)	
	#7	60 TS=(false SAME retest*)	
	#8	267 TS=(retest* SAME negative)	
	#9	25,226 (#1 or #2 or #3 or #4)	
	#10	46,453 (#5 or #6 or #7 or #8)	
	#11	1,344 #9 and #10	
	#12	>100,000 TS=("quality of life")	
	#13	>100,000 TS=(stress* or anxiety or anxious or worry or worried or fright*)	
	#14	>100,000 Ts=(psychol*)	
	#15	48,994 TS=(psychosocial*)	
	#16	>100,000 TS=(depress* or emotion* or fear* or behavior* or behaviour* or confused or confusion)	
	#17	>100,000 (#12 or #13 or #14 or #15 or #16)	
	#18	144 (#11 and #17)	
	#19	>100,000 TS=(qualitative*)	
	#20	>100,000 TS=(evaluat* or analys* or analyt* or ehnograph* or ethnonurs* or validity)	
	#21	>100,000 TS=(intervention*)	
	#22	>100,000 TS=(cluster* or sampl*)	
	#23	>100,000 TS=(interview*)	
	#24	>100,000 TS=(trial* or study or studies)	
	#25	47,396 TS=(phenomolog* or story or stories)	
	#26	60,726 TS=(life experience*)	
	#27	42,972 TS=(focus group*)	
	#28	17,445 TS=(narrative)	
	#29	>100,000 TS=(support* or literature or booklet* or leaflet* or pamphlet* or letter* or video* or podcast* or telephon* or transtelephon*)	
	#30	56,051 TS=(communicat* or counsel*)	
	#31	>100,000 TS=(discuss* or feedback or personalised or personalized or assessment* or reassurance or reassur*)	
	#32	>100,000 Ts=(personal or interpersonall)	
	#33	>100,000 TS=(session* or group* or scheme*)	
	#34	27,100 TS=(patient education)	
	#35	>100,000 (#19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34)	
	#36	123 #18 and #35 - *note all 123 results downloaded*	
	#37	21 #18 not #36 - *note additional 6 selected from here*.	
PsychINFO EBSCO 1887–2010 Searched 6 December 2010	S1	DE "Mammography" (703)	28
	S2	TX (mammogram* or mammograph*) (1268)	
	S3	TX . "National Health Service Breast Screening Programme" (4)	
	S4	TX NHSBSP (2)	
	S5	UK breast screen* program* (8)	
	S6	NHS breast screen* program* (7)	
	S7	S1 or S2 or S3 or S4 or S5 or S6 (1274)	
	S8	TX false* positive* (2109)	
	S9	S7 and S8 (28)	

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)		Number of results
CINAHL EBSCO 1981–2010 Searched 6 December 2010	S1	(MH "Mammography/PF") (146)	99
	S2	(MH "Mammography") (4578)	
	S3	TX (mammogram* or mammograph*) (5651)	
	S4	TX "National Health Service Breast Screening Programme (18)	
	S5	TX NHSBSP (16)	
	S6	TX UK breast screen* program* (21)	
	S7	TX NHS breast screen* program* (43)	
	S8	S2 or S3 or S4 or S5 or S6 or S7 (5683)	
	S9	(MH "False Positive Results") (2001)	
	S10	TX false* positive* or TX "false-positive" (3302)	
	S11	S9 or S10 (3302)	
	S12	S8 and S11 (210)	
	S13	S1 and S11 (20)	
	S14	(MH "Stress+") (31,953)	
	S15	(MH "Psychological Well-Being") OR (MH "Adaptation, Psychological") (14,735)	
	S16	(MH "Anxiety") OR (MH "Anticipatory Anxiety") (10,394)	
	S17	(MH "Psychology+") (6182)	
	S18	(MH "Behavior+") (325,908)	
	S19	(MH "Help Seeking Behavior") (1831)	
	S20	(MH "Quality of Life+") (32,163)	
	S21	(MH "Depression+") (31,871)	
	S22	(MH "Emotions+") (35,643)	
	S23	TX (affect* or anger* or anxiety or anxious or alarm* or attitude* or appetite or behavior* or behaviour* or comply or complian* or coping or concern* or confusion or confused) (382,003)	
	S24	TX (demotivated or demotivation* or de-motivated or de-motivation* or disconcert* or depression or depressed or distress* or deleterious or disappointment or emotion* or fear* or fright* or harm* or mental* or mistrust* or mood* or motivated or motivation* or nightmare* or psychological or psychologically or psychology or psychosocial) (402,584)	
	S25	TX (sleep or stress* or terror or terrified or trust* or mistrust or wellbeing or "well being" or "well-being" or worry or worried or worries) (133,261)	
	S26	S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 (766,002)	
	S27	S12 and S26 (87)	
	S28	S13 or S27 (87)	
	S29	(MH "Rescreening") (52)	
	S30	S2 and S29 (5)	
	S31	S28 or S30 (92)	
	S32	(MH "Qualitative Studies") (33,774)	
	S33	(MH "Clinical Nursing Research") (2266)	
	S34	(MH "Questionnaires+") (124,819)	
	S35	(MH "Attitude+") (152,372)	
	S36	(MH "Focus Groups") (14,223)	
	S37	(MH "Discourse Analysis") (1391)	

Databases, host, date searched, years	Search strategy, keywords added to Reference Manager (Thomson ResearchSoft, San Francisco, CA, USA)	Number of results
	S38 (MH "Content Analysis") (12,487) S39 (MH "Ethnographic Research") (3238) S40 (MH "Ethnonursing Research") (127) S41 (MH "Constant Comparative Method") (3941) S42 (MH "Qualitative Validity") (345) S43 (MH "Purposive Sample") (11,853) S44 (MH "Observational Methods+") (10,984) S45 (MH "Field Studies") (1256) S46 (MH "Theoretical Sample") (921) S47 (MH "Phenomenology") OR (MH "Phenomenological Research") (7951) S48 (MH "Life Experiences+") (10,200) S49 (MH "Cluster Sample+") (1565) S50 S32 or S33 or S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 (284,910) S51 S12 and S50 (27) S52 S13 and S50 (10) S53 S31 or S51 or S52 (94) S54 TX intervention* (123,688) S55 TX program* N5 psycholog* (952) S56 S54 or S55 (124,358) S57 S12 and S56 (25) S58 S13 and S54 (2) S59 S53 or S57 or S58 (99)	
Sociological Abstracts Host: CSA Illumina 1952–2010 Searched 6 December 2010	(mammograph* or mammogram*) and (false* or retest or "re-test") 9 results 1 downloaded	1
IBSS Host: CSA Illumina 1951–2010 Searched 6 December 2010	(mammograph* or mammogram* or (breast scceen*)) and ((false* positive*) or "false-positive")) results Mammography as a descriptor only 3 results. 1. Testing a model of mammography intention Baumann, Linda J.; Brown, R.L.; Fontana, Susan A.; Cameron, Linda Journal of applied social psychology, vol. 23 no. 21, pp. 1733-1756, Nov 1993 View Record | Cited by 4 | Database: IBSS: International Bibliography of the Social Sciences Descriptors: Model testing | Habits | Social influence | Cancer | Health promotion | Women's health | More... 2. Determinants of mammography intentions among prior screenees and nonscreenees Jepson, Christopher; Rimer, Barbara K. Journal of applied social psychology, vol. 23 no. 1, pp. 40-51, Jan 1993 View Record | Cited by 3 | Database: IBSS: International Bibliography of the Social Sciences Descriptors: Health care | Patients | Prevention | Women | Mammography 3. Cancer-related channel selection: a focus on women who have had a mammography Johnson, J. David; Meischke, Hendrika Journal of applied social psychology, vol. 22 no. 24, pp. 1879-1893, Dec 1992 Mammograph* or mammogram* 66 journal results. Nothing relating to false positive results. (mammograph* or mammogram*) and (retest or "re-test" or false test*) 0 results	0

Checked Reference Manager database for these items after PsychLIT search: all added in manually to Reference Manager 9 December 2010.

A model of the influence of false-positive mammography screening results on subsequent screening.
Health Psychol Rev 2010;**4**:112–27.
Authors: DeFranka JT, Brewera N.
Affiliation: a Department of Health Behavior and Health Education, UNC Gillings School of Global Public Health, Chapel Hill, NC, USA.

How possible benefits and harms related to screening for breast cancer are presented to women on web sites.
Barcelona, Spain: XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003. p. 17.
Authors: Jorgensen K, Gotzsche P.
Cochrane Methodology Register (CMR) keywords CMR: Review methodology – applicability and recommendations – individual risk; CMR: other methodology – Internet; CMRA3.
ID CMR-5160.
The Cochrane Methodology Register 2010 Issue 4.

False-negative results in screening programmes: systematic review of impact and implications.
Health Technol Assess 2000;**4**(5).
Authors: Petticrew MP, Sowden AJ, Lister Sharp D, Wright K.
CMR keywords CMR: Evaluation methdology - diagnostic test accuracy - general;CMRA4
ID CMR-3593.

Are benefits and harms in mammography screening given equal attention in scientific articles? A cross-sectional study.
BMC Medicine 2007;**5**:12.
Authors: Jorgensen KJ, Klahn A, Gotzsche PC.
CMR keywords CMR: Other methodology – adverse effects; CMR: Evaluation methodology – bias in trials – outcome reporting bias; CMR: Evaluation methodology – bias in trials – general; CMRA3.
Correspondence address: The Nordic Cochrane Centre, Rigshospitalet, Blegdamsvej 9, Copenhagen, Denmark. kj@cochrane.dk.
ID CMR-10812.
The Cochrane Methodology Register 2010 Issue 4.

Other search strategies undertaken are available from the authors.

Appendix 3 Data extraction forms

Design		Participants		Arms		Outcomes	
Author and year	Brett and Austoker 2001[59]	**Inclusion criteria**	Women invited for routine screening by mammography, already participating in the study at 5 months	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	PCQ, satisfaction with the breast screening service
Study design	Prospective cohort	**Exclusion criteria**	Aged >65 years, symptomatic referral, in another study, developed cancer	***N***	$n = 375$	**Screening attendance**	Intention to reattend and actual reattendance
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	$n = 505$	**Control**	Routine screening by mammography with a normal result		
No. of centres	13			***N***	$n = 130$		
Length of follow-up	35 months						
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign						
Conflicts of interest	None reported						
Notes							
Definition of false-positive	Women, who after attending breast screening units and undergoing further investigations, were not diagnosed with cancer						
Aim	(1) Are women who had a false-positive screening result still having adverse PCs prior to their next routine screen 3 years later? (2) If yes, is the extent of their distress dependent on the processes used in their assessment (e.g. FNA)? (3) If women do experience false-positive adverse psychological effects, does this affect their reattendance?						
	This is the latest publication from a longitudinal study going back to 1995 (see Brett *et al*. 1998[103] and Ong *et al*. 1997[104])						

Methodological issues			
Allocation to groups	NA		
Data analysis	Pearson's chi-squared test for dichotomous data between groups, McNemar's chi-squared test for differences within groups. RRs with CIs were also calculated and Spearman's bivariate correlation for tests of associations between variables. Logistic regression was used to adjust for possible confounding factors. SPSS (SPSS Inc., Chicago, IL, USA) with a two-tailed significance level at $p<0.05$ was used for all calculations		
Handling missing data	Not reported	**Ethics approval**	Yes
Power calculation	Not reported		
Subgroup analysis	Yes		
Demographics			
	n/N	**%**	
Married	305/377	81	
Home owner	330/376	88	
Higher or further education	125/376	33	
Results			
Adverse PCs (PCQ) 1 month before next screening (35 months after last appointment)			
Last breast screening results group (1995)	**% PC score >12 (*n/N*) 1998–9**	**RR (95% CI)**	**Significant difference vs clear after mammography**
Clear after mammography (reference group)	25 (25/99)	Baseline	Baseline
Clear after further mammography and CE	32 (30/93)	1.28 (0.82 to 2.00)	NS
Clear after assessment with FNA	45 (30/66)	1.80 (1.17 to 2.77)	$p = 0.007$
Clear after early recall	46 (46/100)	1.78 (1.19 to 2.66)	$p = 0.002$
Clear after surgical biopsy	52 (11/21)	2.07 (1.22 to 3.52)	$p = 0.014$

Comparison of PCs 1 month after last breast screening appointment and 1 month before the next one

Last breast screening results group (1995)	**% PC score >12 (*n*/*N*) 1995**	**% PC score >12 (*n*/*N*) 1998/9**	**Significant difference**
Clear after mammography (reference group)	26 (26/99)	25 (25/99)	NS
Clear after further mammography and CE	51 (47/93)	32 (30/93)	$p = 0.014$
Clear after assessment with FNA	55 (36/66)	45 (30/66)	$p = 0.015$
Clear after early recall	62 (62/100)	46 (46/100)	$p = 0.034$
Clear after surgical biopsy	71 (15/21)	52 (11/21)	$p = 0.024$

Correlation between PCs at 1 month before returning for next routine breast screening and dissatisfaction with past routine breast screening

Statements about last screening appointment	**False-positive screen**	
	Coefficient	***p*-value**
The amount of time spent for verbal communication at assessment	0.240	0.001
Difficulties with taking in verbal information at breast screening appointment because of anxiety	0.288	0.001
Women's understanding of test result	0.205	0.001
Quality of verbal communication	0.206	0.001
Opportunity to talk to somebody after the breast screening appointment	0.352	0.001
Perceived performance of health workers	0.267	0.001
Verbal communication: chance to say what is on one's mind	0.233	0.001
Amount of information provided in advance	0.179	0.003
Amount of written information	0.279	0.001

Intention to reattend: external factors influencing attitudes and anxiety about attending the next routine breast screening in women with a previous false-positive mammogram					
Item	**% (95% CI)**	***n/N***	**Cause worry (%)**	**Cause worry RR (95% CI)**	***p*-value**
Magazine or newspaper article	29 (24 to 34)	83/288	11	1.18 (1.07 to 1.30)	<0.001
Television programme	25 (20 to 30)	72/288	9	1.13 (1.04 to 1.23)	<0.002
Poster or leaflet	17 (13 to 22)	50/288	_	_	NA
Radio programme	13 (9 to 17)	37/288	_	_	NA
GP attitude to screening	24 (19 to 29)	69/288	_	_	NA
Friend	21 (16 to 26)	60/288	_	_	NA
Family	16 (12 to 20)	(47/288)	_	_	NA
Actual reattendance: Numbers of women attending their next routine screening (3 years)					
Previous false-positive mammography			**Previous normal mammography**		
%		***n/N***	**%**	***n/N***	
85		319/375	92	120/130	

CE, clinical examination; CRC, Cancer Research Campaign; NA, not applicable; NS, not significant; PC, psychological consequence.

Design		Participants		Arms		Outcomes	
Author and year	Brett *et al*. 1998[103]	**Inclusion criteria**	Women invited for routine screening by mammography, already participating in the study at 1 month	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	PCQ
Study design	Prospective cohort	**Exclusion criteria**	Aged >65 years, symptomatic referral, in another study, developed cancer	***N***	Women placed on early recall ($n = 23$); further mammography assessment ($n = 51$); FNA ($n = 41$); biopsy ($n = 45$)	**Screening attendance**	Intention to reattend
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	$n = 284$	**Control**	Routine screening by mammography with a normal result		
No. of centres	12			***N***	$n = 52$		
Length of follow-up	5 months						
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign						
Conflicts of interest	None reported						

Notes	
Definition of false-positive	Women who after attending breast screening units and undergoing further investigations were not diagnosed with cancer
Aim	To find out if (a) women who have a false-positive result after routine screening have adverse psychological consequences 5 months later and (b) if yes, is the extent of their suffering dependent on the process of the further assessment This study is a follow-up from Ong *et al*. 1997[104] and prior to Brett and Austoker 2001.[59] For women on early recall this study was 1 month before their next appointment Sixty-nine (24%) women chose not to return the questionnaire
Methodological issues	
Allocation to groups	NA
Data analysis	The Wilcoxon signed-rank test was used to investigate differences between PCs at 1 and 5 months. The Mann–Whitney *U*-test was used to test for differences between PCs in the different categories of false-positive outcome. Spearman's bivariate correlation tested for associations between PCs and experiences of breast screening. Logistic regression was used to explore variables relating to women's breast screening experience. SPSS with a two-tailed significance level at $p<0.05$ was used for all calculations
Handling missing data	Not reported
Ethics approval	Not reported
Power calculation	Sample size based on responders to phase 1 of the study
Subgroup analysis	Yes
Demographics	
Not reported	

Results

Adverse PCs (PCQ) 5 months after their last screening appointment

False-positive subgroup	% PC (*n*/*N*)	Significant difference vs routine recall after mammography	RR (95% CI)
NR after screening	10 (5/52)	Baseline	
NR after assessment without FNA	45 (23/51)	$p<0.0001$	4.7 (1.93 to 11.38)
NR after assessment with FNA	44 (18/41)	$p<0.0001$	4.6 (1.85 to 11.26)
NR after benign biopsy	60 (27/45)	$p<0.00001$	5.11 (2.13 to 12.26)
Early recall (6 months)	61 (14/23)	$p<0.00001$	6.33 (2.59 to 15.50)

Comparison of adverse PCs 1 month and 5 months after last breast screening appointment

False-positive subgroup	% PC (*n*/*N*) 1 month after last appointment	% PC (*n*/*N*) 5 months after last appointment	Significant difference
NR after screening	17 (9/52)	10 (5/52)	NS
NR after assessment without FNA	57 (29/51)	45 (23/51)	$p<0.001$
NR after assessment with FNA	63 (26/41)	44 (18/41)	$p<0.001$
NR after benign biopsy	91 (21/23)	61 (14/23)	$p<0.001$
Early recall (6 months)	70 (32/46)	59 (27/46)	NS

Logistic regression: variables related to PCs at 5 months after the last breast screening appointment

Variable	OR	95% CI	Significance
PCs at 1 month	5.82	2.70 to 12.56	$p<0.001$
Age of women	1.00	0.98 to 1.03	NS
Result group (type of investigation)	4.4	1.35 to 14.35	$p<0.01$
Likelihood of attending future breast screening	0.61	0.03 to 11.93	NS
Greater perceived likelihood of ever getting breast cancer compared with the average woman	0.91	0.35 to 2.34	NS
Apprehensiveness about attending	0.92	0.80 to 1.07	NS
Need to discuss breast screening with someone	0.5	0.24 to 1.02	NS

CRC, Cancer Research Campaign; NA, not applicable; NR, normal recall (3 years); NS, not significant; PC, psychological consequence.

Design		Participants		Arms		Outcomes	
Author and year	Ong *et al*. 1997[104]	**Inclusion criteria**	Women invited for routine screening by mammography who were recalled for assessment	**Intervention**	Women placed on ER (<3 years), *n* = 182	**Psychological**	PCQ
Study design	Cross section	**Exclusion criteria**	Not reported	***N***	*n* = 182	**Screening attendance**	_
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	*n* = 877	**Control**	Women placed on RR		
No. of centres	13			***N***	RR: after mammography (*n* = 173); further mammography assessment (*n* = 166); FNA (*n* = 109); biopsy (*n* = 31)		
Length of follow-up	Measures taken 1 month after assessment			**Notes**			
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign, NHSBSP						
Conflicts of interest	None reported						
Notes							
Definition of false-positive	Not defined						
Aim	To find out if women suffered adverse psychological consequences from being put on ER following a false-positive mammogram and to suggest solutions to reduce them This study was primarily about the effects of early recall on women who had been called back for assessment after their mammogram						

Methodological issues			
Allocation to groups	NA		
Data analysis	Differences between groups were calculated with chi-squared tests, bivariate testing, logistic and multivariate linear regression were used to calculate the influence of single PCQ variables. SPSS with a two-tailed significance level at $p<0.05$ was used for all calculations		
Handling missing data	Median scores were used per item on the PCQ. Those not responding to any items were coded as missing values and excluded from the analysis	**Ethics approval**	Not reported
Power calculation	Yes		
Subgroup analysis	Yes		
Demographics			
Not reported			
Results			
Adverse PCs (PCQ) 1 month after further assessment			

Outcome of last screening visit	**% reporting adverse PCs**	***n/N***	**Significance compared with women placed on RR after mammography**	**Significance compared with women placed on RR after assessment**
RR after mammography	29	38/130		
RR after assessment	50	64/128	$p<0.0005$	
RR after FNA	58	61/106	$p<0.00001$	NS
ER after assessment	63	81/130	$p<0.00001$	$p<0.05$
RR after biopsy	87	26/30	$p<0.00001$	$p<0.0005$

CRC, Cancer Research Campaign; ER, early recall; NA, not applicable; PC, psychological consequences; RR, routine recall.

Design		Participants		Arms		Outcomes	
Author and year	Sutton *et al*. 1995[55]	**Inclusion criteria**	Women invited for routine screening by mammography who were recalled for assessment	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	Ad hoc anxiety questionnaire with a three-point scale
Study design	Retrospective cohort	**Exclusion criteria**	None reported	***N***	*N* = 24	**Screening attendance**	_
Study centre	Institute of Psychiatry, London	***N***	*N* = 1021	**Control**	Routine screening by mammography with a normal result		
No. of centres	1			***N***	*N* = 671		
Length of follow up	9 months after pre-screening baseline						
Setting	NHSBSP mobile screening unit						
Funding	Imperial Cancer Research Fund						
Conflicts of interest	None reported						
Notes							
Definition of false-positive	Women who are recalled for investigation after a positive breast screen but subsequently receive a normal result						
Aim	To find out if mammography raises anxiety in routinely screened women who have a negative result						
Methodological issues							
Allocation to groups	NA						
Data analysis	These included product-moment correlations, independent and paired *t*-tests and repeated measures ANOVA. Only unadjusted results are reported. SPSS with two-tailed significance at 0.05 was used for all calculations						
Handling missing data	Not reported	**Ethics approval**	Not reported				
Power calculation	Not reported						
Subgroup analysis	Yes						

Demographics						
Measured for the whole sample but data only reported for approximately 40% of sample. It is unknown who these 40% were						
Results						
Retrospective anxiety at 9 months after baseline pre-screening: three-point scale (1 = not anxious, 2 = a bit anxious, 3 = very anxious)						
Outcome of last screening visit	**Stage 1: receive screening invitation, mean (SD)**	**Stage 2: while waiting for the mammogram, mean (SD)**	**Stage 3: at the clinic after the mammogram, mean (SD)**	**Stage 4: after screening and before receiving the results, mean (SD)**	**Stage 5: after reading the results letter, mean (SD)**	**Stage 6: now (9 months after baseline), mean (SD)**
False-positive	Not reported	Not reported	1.60 (0.68)	1.95 (0.09)	2.85 (0.37)	Not reported
Normal mammogram	Not reported	Not reported	1.36 (0.52)	1.70 (0.57)	1.16 (0.36)	Not reported
Statistical significance of the difference between the groups			$p < 0.05$	$p = 0.054$	$p < 0.001$	

ANOVA, analysis of variance; NA, not applicable.

Comment: only some of the results were reported numerically. Other scores were reported graphically in such a way that it is difficult to accurately read the scores.

Design		Participants		Arms		Outcomes	
Author and year	Bull and Campbell 1991[106]	**Inclusion criteria**	Women invited for routine screening by mammography who were recalled for assessment	**Intervention**	Women invited for routine screening by mammography who were recalled for assessment	**Psychological**	Ad hoc questionnaire including frequency of breast self-examination HADS
Study design	Prospective cohort	**Exclusion criteria**	Not reported	***N***	Group A: invitation ($n = 541$); group B: normal mammogram ($n = 331$); group C: assessment with mammogram, ultrasound, FNA ($n = 204$); group D: assessment with surgical biopsy ($n = 49$)	**Screening attendance**	–
Study centre	Salisbury and Southampton Health District	***N***	$n = 541$	**Control**	NA		
No. of centres	1			***N***	–		
Length of follow-up	Measures taken 6 weeks after the 'all-clear'						
Setting	Salisbury and Southampton Health District mammography screening programme						
Funding	Not reported						
Conflicts of interest	Not reported						

Notes				
Definition of false-positive	Not reported			
Aim	To assess the psychological effects on well women of participating in the screening programme It is not known if the women had previously had cancer or were in a high-risk group			
Methodological issues				
Allocation to groups	NA			
Data analysis	A paired comparison of women in groups A and B used a paired *t*-test or Wilcoxon rank-sum test. Independent groups were compared using ANOVA or Kruskal–Wallis test			
Handling missing data	Not reported	**Ethics approval**	Not reported	
Power calculation	Yes			
Subgroup analysis	No			
Demographics				
Age (years)	**Group A, *n* (%)**	**Group B, *n* (%)**	**Group C, *n* (%)**	**Group D, *n* (%)**
50–54	122 (22.6)	76 (22.9)	66 (32.3)	10 (20.4)
55–59	154 (28.5)	113 (34.1)	54 (26.5)	18 (36.7)
60–64	185 (34.2)	105 (31.7)	54 (26.5)	16 (32.7)
65–70	40 (7.4)	26 (7.9)	15 (7.4)	4 (8.2)

Results

Frequency of breast self-examination by group	**Group A invite to screening, *n* (%)**	**Group B normal mammogram, *n* (%)**	**Group C false-positive (not biopsy), *n* (%)**	**Group D false-positive (biopsy), *n* (%)**	
Never	56 (18)	22 (22)	24 (12)	7 (14)	
Less than once a month	155 (50)	23 (23)	34 (17)	7 (14)	
Once a month	69 (22)	47 (46)	97 (48)	18 (37)	
Once a week	25 (8)	10 (10)	41 (20)	12 (25)	
More than once a week	6 (2)	0	8 (4)	5 (10)	
No response	1 (0)	0	0	0	
HADS	**Group A**	**Group B**	**Group C**	**Group D**	***p*-value**
Depression scale, mean (range)	5.0 (0–19)	4.23 (0–15)	4.25 (0–16)	3.82 (0–18)	0.0003
Anxiety scale, mean (range)	4.97 (0–20)	4.43 (0–17)	4.32 (0–15)	4.27 (0–14)	0.014
HADS severity of score by group	**Group A invite to screening, *n* (%)**	**Group B normal mammogram, *n* (%)**	**Group C false-positive (not biopsy), *n* (%)**	**Group D false-positive (biopsy), *n* (%)**	***p*-value**
Depression					
Normal (0–7)	232 (75)	95 (91)	168 (83)	43 (88)	NS
Borderline (8–10)	52 (17)	7 (7)	25 (12)	3 (6)	NS
Abnormal (>10)	26 (8)	2 (2)	9 (4)	3 (6)	NS
Anxiety					
Normal (0–7)	253 (81)	91 (88)	174 (86)	42 (86)	NS
Borderline (8–10)	40 (13)	10 (10)	24 (12)	4 (8)	NS
Abnormal (>10)	20 (6)	2 (2)	4 (2)	3 (6)	NS

ANOVA, analysis of variance; NA, not applicable; NS, not significant.

Design		Participants		Arms		Outcomes	
Author and year	Ellman *et al.* 1989[105]	**Inclusion criteria**	Women invited for routine mammography screening, those recalled for further assessment and those with symptoms being further investigated	**Intervention**	Group B: routine screening by mammography with a false-positive result	**Psychological**	GHQ-28, ad hoc questionnaire
Study design	Prospective cohort	**Exclusion criteria**	Not reported	***N***	$n = 271$	**Screening attendance**	–
Study centre	Institute of Cancer Research, Sutton, Surrey	***N***	$n = 752$	**Control**	Routine screening by mammography with a normal result, symptomatic women who did not have cancer, symptomatic or recalled screened women who did have cancer, history of breast cancer with or without symptoms		
No. of centres	1			***N***	Group A: routine screening by mammography with a normal result ($n = 295$); *group C: symptomatic women who did not have cancer* ($n = 134$); *group D: symptomatic or recalled screened women who did have cancer* ($n = 38$); *group E: history of breast cancer with or without symptoms* ($n = 14$)		
Length of follow-up	3 months after clinic attendance			**Notes**	Participants also received clinical examination. *Symptomatic women do not meet the inclusion criteria for this review and are not included. Those with a history of breast cancer are also excluded in this case because those with and without symptoms were aggregated*		

Design	Participants	Arms	Outcomes
Setting	South West Surrey Health District breast screening programme		
Funding	DHSS Research Management Division		
Conflicts of interest	None reported		
Notes			
Definition of false-positive	Women who attended breast cancer screening clinics who were recalled for further investigation which showed no cancer		
Aim	To find out the immediate and persistent psychiatric morbidity in women recalled for further assessment following mammography screening		
Methodological issues			
Allocation to groups	NA		
Data analysis	Groups' scores were compared with a score of at least 5 to indicate probable psychiatric morbidity, using chi-squared. Change scores were analysed with the Wilcoxon test and between-groups' scores with the Mann–Whitney *U*-test. All tests were two-tailed with significance at $p<0.05$		
Handling missing data	Not reported	**Ethics approval**	Not reported
Power calculation	Not reported		
Subgroup analysis	No		
Participant characteristics			
	Group A normal mammogram	**Group B false-positive**	
Total recruited	295	271	
No. completing both questionnaires (%)	287 (97.3)	266 (98.2)	
Mean age (+– SD)	53.9 (6.8)	54.5 (7.4)	
First screening %	18.3	20.7	

Results

Proportion of GHQ scores of at least 5 at the screening clinic and 3 months later

	Group A normal mammogram	**Group B false-positive**	***p*-value**
Screening visit (95% CI)	24.0% (20% to 30%)	30.1% (24% to 36%)	NS
3 months later (95% CI)	19.2% (15% to 24%)	18.8% (14% to 24%)	NS

Distribution of GHQ scores

	Group A normal mammogram, no. (%)	**Group B false-positive, no. (%)**
Screening visit score		
0	118 (40.0)	111 (41.0)
1–4	104 (35.3)	78 (28.8)
5–9	49 (16.6)	48 (17.7)
10–28	24 (8.1)	34 (12.5)
Total	295 (100)	271 (100)
3 months later score		
0	150 (52.3)	157 (59.0)
1–4	82 (28.6)	59 (22.2)
5–9	31 (10.8)	23 (8.6)
10–28	24 (8.4)	27 (10.2)
Total	287 (100)	266 (100)

Distribution of GHQ subscale scores				
Symptom subscale	**Group A normal mammogram, no. (%)**		**Group B false-positive, no. (%)**	
	Screening visit (*n* = 295)	**3 months later (*n* = 287)**	**Screening visit (*n* = 271)**	**3 months later (*n* = 266)**
Somatic	113 (38)	98 (34)	108 (40)	69 (26)
Anxiety	104 (35)	75 (26)	119 (44)	77 (29)
Social dysfunction	104 (35)	86 (30)	89 (33)	77 (29)
Depression	42 (14)	29 (10)	38 (14)	27 (10)
Ad hoc questionnaire: opinions about the breast screening clinic				
	Groups A and B			
Criticism of communication	40 (7%)			

NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Brain *et al.* 2008[102]	**Inclusion criteria**	Women aged 35–49 years invited for routine annual screening by mammography with a FHBC	**Intervention**	Routine annual screening by mammography with a false-positive result	**Psychological**	Questionnaire including: CWS-R, cognitive appraisal, brief COPE, perceived risk of breast cancer, dispositional optimism
Study design	Prospective cohort	**Exclusion criteria**	Previous history of breast cancer or family history of ovarian cancer	***N***	$n = 112$	**Screening attendance**	_
Study centre	Institute of Medical Genetics, University of Cardiff PIMMS Management Group	***N***	$n = 1250$	**Control**	Routine annual screening by mammography with a normal result		
No. of centres	21			***N***	$n = 1174$		
Length of follow-up	6 months' measures taken at T1 1 month before screening, T2 1 month, and T3 6 months after the 'all-clear'						
Setting	NHS screening clinics for women with FHBC						
Funding	Cancer Research UK						
Conflicts of interest	Not reported						

Notes			
Definition of false-positive	Women who attend screening and are recalled for further investigations before being given the 'all-clear'		
Aim	This study aimed to find pre-screening variables that predicted cancer-specific distress 1 and 6 months after screening		
Methodological issues			
Allocation to groups	NA		
Data analysis	Changes in scores were compared with paired *t*-tests. Preliminary associations were tested with partial correlations. Hierarchical multiple regression explored the contributions of independent variables		
Handling missing data	Not reported	**Ethics approval**	Yes
Power calculation	In related paper Tyndel *et al*. 2007[101]		
Subgroup analysis	No		
Demographics			
Participant characteristics from Tyndel et al. *2007*[101]			
Item	**Recall result (*n* = 112)**	**Normal result (*n* = 1174)**	
	No. (%)	**No. (%)**	
Age, mean (SD)	43.2 (3.52)	43.2 (3.44)	
Ethnic group – white	109 (97.3)	1157 (98.6)	
Married or partner	109 (97.3)	1158 (98.6)	
Higher education	108 (96.4)	1155 (98.3)	
Have biological children	109 (97.3)	1158 (98.6)	

Results

Multiple regression showing predictive associations between independent baseline variables and cancer worry scores at 1 and 6 months

T1 variable (1 month before screening)	T2 (1 month after screening) CWS-R	*p*-value	T3 (6 months after screening) CWS-R	*p*-value
T1 cancer worry	0.543	<0.001	0.581	<0.001
High perceived lifetime risk of breast cancer	0.092	<0.001	0.075	<0.01
Relative died of breast cancer in the last year	–	–	0.050	<0.05
Belief in increased risk due to family history	0.091	<0.001	0.082	<0.001
First attendance at the screening programme	–0.067	<0.001	–0.044	<0.05
Being recalled for further tests	0.061	<0.05	–	–
Low emotion focused coping potential	–0.055	<0.05	–0.053	<0.05
Use of religion as a coping strategy	0.050	<0.01	–	–
Dispositional optimism	–0.045	<0.05	–0.003	NS
Low challenge appraisal	–0.043	<0.05	–0.019	NS
Substance use for coping	0.042	<0.05	–	–

NA, not applicable.

Design		Participants		Outcomes	
Author and year	Clements *et al*. 2008[107]	**Inclusion criteria**	Women aged 35–50 years invited for routine annual screening by mammography with a FHBC	**Psychological**	The value women placed on being on a FHBC annual screening programme and their reactions to either having an initial all-clear result after screening or only have this result after further investigation (false-positive)
Study design	Interview	**Exclusion criteria**	Previous history of breast cancer or family history of ovarian cancer		
Theoretical framework	Not reported	***N***	$n = 58$: normal result $n = 36$; false-positive $n = 22$		
Study centre	Primary Care Education Research Group, University of Oxford PIMMS Management Group				
Time from 'all-clear'	Not reported				
Setting	NHS screening clinics for women with FHBC				
Funding	Cancer Research UK				
Conflicts of interest	Not reported				

Notes	
This research has only been published as a summary of a poster. It is only included because it is a nested study in Tyndel *et al*. 2007[101]	
Definition of false-positive	Women who were recalled for further tests prior to an all-clear result
Aim	To explore the value women placed on being part of a screening programme and to understand the reactions of women who had false-positive results
Methodological issues	
Sampling strategy	Women who were participants in the Tyndel *et al*. 2007 study[101]
Data analysis	Thematic
All a priori outcomes reported	Yes
Demographics	
Note reported	
Results	
These were only briefly summarised: Women believed that participating in screening would enable cancer to be detected at an early stage leading to a positive outcome Women had greater faith in mammography than themselves to detect early cancer An all-clear result gave a high degree of reassurance that they did not have cancer Women with a false-positive result were initially distressed, the all-clear gave increased feelings of reassurance and security and a greater faith in screening than those with an initial all-clear result Being recalled was given a positive interpretation as proof that screening worked Fear of breast cancer was relieved by being part of the breast screening programme and made the women feel more in control of their family history	

Design		Participants		Arms		Outcomes	
Author and year	Tyndel *et al*. 2007[101]	**Inclusion criteria**	Women aged 35–49 years invited for routine annual screening by mammography with a FHBC	**Intervention**	Routine annual screening by mammography with a false-positive result	**Psychological**	CWS-R, PCQ
Study design	Prospective cohort	**Exclusion criteria**	Previous history of breast cancer or family history of ovarian cancer	***N***	$n = 166$	**Screening attendance**	_
Study centre	Primary Care Education Research Group, University of Oxford PIMMS Management Group			**Control**	Routine annual screening by mammography with a normal result		
No. of centres	21			***N***	$n = 2084$		
Length of follow-up	6 months measures taken at 1 month before screening and 1 and 6 months after the 'all-clear'						
Setting	NHS screening clinics for women with FHBC						
Funding	Cancer Research UK						
Conflicts of interest	None						

Notes				
Definition of false-positive	Not reported			
Aim	To test the hypothesis that in the short and long term women who receive an immediate all-clear result gain psychological benefit from screening, whereas women who are recalled for additional tests before an all-clear result experience increased cancer-specific distress			
Methodological issues				
Allocation to groups	NA			
Data analysis	Between-group categorical characteristics were compared with chi-squared and continuous variables with the Mann–Whitney *U*-test. Negative psychological effects at follow-up were analysed with linear regression with a preliminary analysis using the Mann–Whitney *U*-test			
Handling missing data	Not reported	**Ethics approval**	Yes	
Power calculation	Yes			
Subgroup analysis	No			
Demographics				
Participant characteristics				
Item	**Recall result (*n* = 112)**	**Normal result (*N* = 1174)**	**Mann–Whitney *U*-test and chi-squared test**	***p*-value**
	No. (%)	**No. (%)**		
Age, mean (SD)	43.2 (3.52)	43.2 (3.44)	65,468	NS
Ethnic group – white	109 (97.3)	1157 (98.6)	1.595	NS
Married or partner	109 (97.3)	1158 (98.6)	0.018	NS
Higher education	108 (96.4)	1155 (98.3)	0.305	NS
Have biological children	109 (97.3)	1158 (98.6)	4.896	0.027
High familial risk	109 (97.3)	1166 (99.3)	4.417	0.036
Hospital attendance for recall assessment	112 (100)	1167 (99.4)	56.850	0.000

Results

Within group comparison of distress at T1 (1 month before screening), T2 (1 month after screening) and T3 (6 months after screening)

Questionnaire	False-positive result	Within false-positive result		Normal result	Within normal result	
	Mean (SD)	Paired *t*-test	*p*-value	Mean (SD)	Paired *t*-test	*p*-value
CWS-R						
T1 ($n = 111, 1171$)	11.61 (2.90)	–	–	10.99 (2.91)	–	–
T2 ($n = 111, 1171$)	11.68 (2.89)	–	–	10.56 (2.60)	–	–
T3 ($n = 111, 1159$)	10.35 (2.65)	–	–	10.12 (2.49)	–	–
Difference T1–T2		–0.298	NS	-	7.537	<0.01
Difference T2–T3		6.372	<0.01	-	8.633	<0.01
PCQ						
T1 ($n = 110, 1167$)	7.32 (7.66)	–	–	5.06 (6.71)	–	–
T2 ($n = 110, 1167$)	7.1 (7.44)	–	–	4.18 (6.19)	–	–
T3 ($n = 110, 1169$)	4.61 (6.42)	–	–	3.84 (6.00)	–	–
Difference T1–T2	–	–0.051	NS	–	6.935	<0.01
Difference T2–T3	–	5.752	<0.01	–	3.183	<0.01

Between-group impact of false-positive result on positive outcomes at T2 and T3

Outcome	False-positive result	Normal result	Mann–Whitney *U*-test	95% CI	*p*-value
Positive PCQ at T2	–	–	51,561		0.002
Mean (SD)	13.02 (7.6)	10.81 (6.9)	–	–	–
Positive PCQ at T3	–	–	–	–	–
Mean (SD)	12.65 (8.9)	11.16 (7.0)	59,169		NS
			OR		
Benefits of screening more positive at T2	–	–	3.168	2.138 to 4.696	0.00
No. (%)	112 (55)	1164 (27)	–	–	–
Benefits of screening more positive at T3					
No. (%)	105 (35)	1085 (19)	2.35	1.531 to 3.606	0.00

NA, not applicable; NS, not significant.

Design		Participants		Arms		Outcomes	
Author and year	McCann *et al.* 2002[61]	**Inclusion criteria**	Women aged 49–63 years invited for routine breast screening by mammography	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	_
Study design	Retrospective cohort	**Exclusion criteria**	Women who were aged > 63 years at follow-up	***N***	$n = 4792$	**Screening attendance**	Subsequent attendance at routine screening after a false-positive result and rate of interval cancer – from records
Study centre	Cancer Intelligence Unit, University of Cambridge	***N***	$n = 140,387$	**Control**	Routine screening by mammography with a normal result	**Quality of life**	_
No. of centres	Not reported			***N***	$n = 108,617$		
Length of follow-up	3.5 years						
Setting	NHSBSP in East Anglia						
Funding	NHS Executive Eastern Region						
Conflicts of interest	Not reported						

Notes			
Definition of false-positive	Any woman who is recalled for assessment on the basis on mammographic findings and in whom cancer is not diagnosed		
Aim	To find out if false-positive mammography affects reattendance in East Anglia, to quantify the increased risk of interval cancer and to determine if the risk of cancer detection at second screening is increased		
Methodological issues			
Allocation to groups	NA		
Data analysis	Not reported		
Handling missing data	Not reported	**Ethics approval**	NA
Power calculation	NA		
Subgroup analysis	Yes		
Demographics			
	False-positive, mean (SD)	**Normal, mean (SD)**	
Age	56.1 (3.5)	55.8 (3.5)	

Results

Likelihood of reattendance at second round cancer screening (3 years later)

Study group	*n* (%)	95% CI	OR (95% CI)
All groups			
All	97,062 (85.6)	85.4 to 85.8	
With interval cancer	72 (19.2)	15.2 to 23.2	
Without interval cancer	96,990 (85.8)	85.6 to 86.0	
Normal result			
All	93,081 (85.7)	85.5 to 85.9	1
With interval cancer	69 (21.0)	16.6 to 25.4	
Without interval cancer	93,012 (85.9)	68.1 to 85.7	
False-positive – no biopsy			
All	3572 (83.5)	82.4 to 84.6	0.84 (0.78 to 0.92)
With interval cancer	3 (7.1)	0 to 14.9	
Without interval cancer	3569 (84.3)	83.2 to 85.4	
False-positive – biopsy			
All	409 (79.6)	76.1 to 83.1	0.65 (0.52 to 0.81)
With interval cancer	0	0	
Without interval cancer	409 (80.2)	76.7 to 83.7	
False-positive – all			
All	3981 (83.1)	82.0 to 84.4	0.82 (0.76 to 0.89)
With interval cancer	3 (6.5)	0 to 13.7	
Without interval cancer	3978 (83.8)	82.8 to 84.9	

NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	O'Sullivan *et al.* 2001[108]	**Inclusion criteria**	Women invited for mammography screening for the second or subsequent time	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	_
Study design	Retrospective cohort	**Exclusion criteria**	Women invited for the first time and women who had been previously invited but had never attended	***N***	$n = 248$	**Screening attendance**	Subsequent attendance at routine screening after a false-positive result – from records
Study centre	Department of Psychology, University of Essex	***N***	$n = 5649$	**Control**	Routine screening by mammography with a normal result	**Quality of life**	_
No. of centres	Not reported			***N***	$n = 5401$		
Length of follow-up	Not reported						
Setting	East London and City of London Health Districts						
Funding	Cancer Research Campaign						
Conflicts of interest	Not reported						
Notes							
Definition of false-positive	Women who have previously experienced an abnormal breast screening result, which after further assessment was concluded to be negative for malignancy						
Aim	Effects of a false-positive result on reattendance for those on early recall and routine recall						

Methodological issues			
Allocation to groups	NA		
Data analysis	Not reported		
Handling missing data	NA	**Ethics approval**	NA
Power calculation	NA		
Subgroup analysis	No		
Demographics			
Not reported			
Results			
Attendance at second screening			
Result at initial screening	**Attend second screen, *N* (%)**	**Do not attend second screen, *N* (%)**	**Total**
Normal	3841 (71)	1560 (29)	5401
False-positive – all	175 (70.6)	73 (29.4)	248
False-positive – routine recall	119 (73.5)	43 (26.5)	162
False-positive – early recall	56 (65)	30 (35)	86
Total	4016	1633	5649

NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Meldrum *et al.* 1994[115]	**No. randomised**	3083	**Intervention**	Tailored invitation accounting for screening history for second round mammography screening	**Psychological**	–
Study design	RCT	**Inclusion criteria**	All women invited for second round routine mammography screening	***N***	False-positive $n = 115$; normal $n = 800$	**Screening attendance**	Subsequent attendance at routine screening and effect of a tailored invitation on subgroups
Study centre	Department Public Health, Glasgow Royal Maternity Hospital	**Exclusion criteria**	Women with breast cancer and those whose screening history was not available	**Control**	Standard invitation for second round mammography screening	**Quality of life**	–
No. of centres	1			***N***	False-positive; $n = 112$; normal $n = 791$		
Length of follow-up	Not reported						
Registered	Pre-dates registration						
Setting	North West Glasgow Breast Screening Centre						
Funding	Scottish Office Home and Health Department						
Conflicts of interest	Not reported						

Notes	
Definition of false-positive	Women who attended and were recalled for further tests before they were given an all-clear result
Aim	To determine if attendance at second-round screening (3 years later) could be improved by the use of invitation letters tailored to the outcome of the previous screening round
Methodological issues	
Randomisation and allocation	Random number tables were used. Randomisation was within-study group (false-positive or normal) to intervention or control. It is unclear if the participants were aware that they were in a trial. It appears that they were sent one of two letters from the screening centre; it is unclear whether or not the assessors knew which group women were in
Data analysis	Between-group differences were tested by chi-squared. Analysis was by intention to treat
Missing data	Not reported
Power calculation	Yes
Subgroup analysis	No
All a priori outcomes reported	Unknown
Baseline characteristics	
Not reported	

Results

Second-round screening attendance: comparing standard vs tailored letters within groups

	Previously false-positive			Previously all clear		
	Standard letter	**Tailored letter**	**% difference (95% CI), *p*-value**	**Standard letter**	**Tailored letter**	**% difference (95% CI), *p*-value**
Invited, *N*	112	115		791	800	
Attended, *N*	78	94		583	594	
Attended, % (95% CI)	70 (61 to 78)	82 (75 to 89)	12.1 (1.03 to 23.2), 0.03	74 (71 to 77)	74 (71 to 77)	0.5 (–3.8 to 4.9), 0.8

Second-round screening attendance: comparing standard vs tailored letters between groups

	Standard letter			Tailored letter		
	Previously false-positive	**Previously all clear**	**% difference (95% CI)**	**Previously false-positive**	**Previously all clear**	**% difference (95% CI)**
Invited, *N*	112	791		115	800	
Attended, *N*	78	583		94	594	
Attended, % (95% CI)	70 (61 to 78)	74 (71 to 77)	–0.04 (–0.13 to 0.05)	82 (75 to 89)	74 (71 to 77)	0.08 (0.003 to 0.157)

Meldrum *et al.* 1994[115]

Section/topic	Item No.	Compliant	Checklist item
Title and abstract			
	1a	Yes	Identification as a randomised trial in the title
	1b	Yes	Structured summary of trial design, methods, results and conclusions (for specific guidance see CONSORT for abstracts)
Introduction			
Background and objectives	2a	Yes	Scientific background and explanation of rationale
	2b	Yes	Specific objectives or hypotheses
Methods			
Trial design	3a	Not reported	Description of trial design (such as parallel, factorial) including allocation ratio
	3b	NA	Important changes to methods after trial commencement (such as eligibility criteria), with reasons
Participants	4a	Yes	Eligibility criteria for participants
	4b	Yes	Settings and locations where the data were collected
Interventions	5	No	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Yes	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
	6b	NA	Any changes to trial outcomes after the trial commenced, with reasons
Sample size	7a	Yes	How sample size was determined
	7b	NA	When applicable, explanation of any interim analyses and stopping guidelines
Randomisation:			
Sequence generation	8a	Yes	Method used to generate the random allocation sequence
	8b	Not reported	Type of randomisation; details of any restriction (such as blocking and block size)
Allocation concealment mechanism	9	Not reported	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned
Implementation	10	Not reported	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions
Blinding	11a	Not reported	If done, who was blinded after assignment to interventions (e.g. participants, care providers, those assessing outcomes) and how
	11b	Yes	If relevant, description of the similarity of interventions

Section/topic	Item No.	Compliant	Checklist item
Statistical methods	12a	Yes	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Yes	Methods for additional analyses, such as subgroup analyses and adjusted analyses
Results			
Participant flow (a diagram is strongly recommended)	13a	Yes	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	No	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Yes	Dates defining the periods of recruitment and follow-up
	14b	NA	Why the trial ended or was stopped
Baseline data	15	No	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	Yes	For each group, number of participants (denominator) included in each analysis and whether or not the analysis was by original assigned groups
Outcomes and estimation	17a	Yes	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% CI)
	17b	NA	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	NA	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory
Harms	19	Not reported	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)
Discussion			
Limitations	20	Not reported	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses
Generalisability	21	Not reported	Generalisability (external validity, applicability) of the trial findings
Interpretation	22	No	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Other information			
Registration	23	Pre-registry	Registration number and name of trial registry
Protocol	24	No	Where the full trial protocol can be accessed, if available
Funding	25	Yes	Sources of funding and other support (such as supply of drugs), role of funders

NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Orton *et al*. 1991[109]	**Inclusion criteria**	Women aged 45–64 invited to attend for second-round screening by mammography	**Intervention**	Routine screening by mammography with a false-positive result	**Psychological**	Acceptability of screening
Study design	Cross section	**Exclusion criteria**	Not reported	***N***	$n = 50$	**Screening attendance**	Reattendance
Study centre	Aylesbury, Oxfordshire	***N***	$n = 1582$	**Control**	Routine screening by mammography with a normal result	**Quality of life**	_
No. of centres	1			***N***	$n = 1532$		
Length of follow-up	NA						
Setting	Breast screening in Aylesbury Vale						
Funding	Not reported						
Conflicts of interest	Not reported						

Notes			
Definition of false-positive	If after screening a woman is asked to reattend for further assessment but no malignancy is found		
Aim	To find out whether the acceptability of screening or having a false-positive mammogram affected attendance at subsequent breast screening Only the measure of reattendance was disaggregated and is reported		
Methodological Issues			
Allocation to groups	NA		
Data analysis	Data were analysed with a chi-squared test		
Handling missing data	Not reported	**Ethics approval**	Not reported
Power calculation	Not reported		
Subgroup analysis	No		
Demographics			
Not reported			
Results			
Attendance at second-round screening			
	False-positive, *N* (%)	**Normal result, *N* (%)**	
Invited	50 (100)	1532 (100)	
Attended	46 (92)	1362 (89)	

NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Ong and Austoker 1997[110]	**Inclusion criteria**	Women invited for routine screening by mammography who were recalled for assessment	**Intervention**	Women invited for routine screening by mammography who were recalled for assessment	**Psychological**	Ad hoc questionnaire about the acceptability of information given in assessment invitations
Study design	Cross section	**Exclusion criteria**	Women recalled due to poor quality X-rays	***N***	*n* = 1493	**Screening attendance**	_
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	*n* = 1493	**Control**	NA	**Quality of life**	_
No. of centres	8			***N***	NA		
Length of follow-up	NA						
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign and the NHSBSP						
Conflicts of interest	Not reported						
Notes							
Definition of false-positive	Not reported						
Aim	Evaluation of women's experiences at the assessment clinic and their information needs there and afterwards, including a discourse analysis of open questions						

Methodological issues			
Allocation to groups	NA		
Data analysis	Contingency tables were used for comparison		
Handling missing data	Not reported	**Ethics approval**	Yes
Power calculation	Not reported		
Subgroup analysis	Yes/no		
Demographics			
Not reported			
Results			
Communication at the assessment centre and level of distress			

Communication	**Distressed/very distressed, % (*n*/*N*)**	**Somewhat/not distressed, % (*n*/*N*)**	***p*-value**
Women who had not talked with somebody at the centre about the reason for recall	33 (275/835)	32 (191/597)	NS
Women who would have liked to talk about the reason for recall	26 (214/835)	18 (108/597)	<0.0001
Women who thought they were not given enough information about the physical examination they had	6 (46/757)	4 (20/563)	<0.05
Women who thought they were not given enough information about the X-rays they had	9 (72/773)	4 (22/553)	<0.0005

Communication at the assessment centre and the role of breast care nurses

Communication	Centres where women were not systematically provided with the opportunity to talk immediately before tests, % (*n*/*N*)	Centres where the breast care nurse provided women with the opportunity to talk in private immediately before tests, % (*n*/*N*)	*p*-value
Women who had talked at the centre about reason for recall:			
With 'somebody at the centre'	58 (611/1055)	93 (374/401)	<0.001
With a doctor or radiologist	31 (323/1035)	7 (26/391)	<0.001
With a nurse	9 (97/1035)	60 (234/391)	<0.001
Women who would have liked to talk about reason for recall	30 (310/1039)	4 (16/400)	<0.001
Women who stated that the test they had were not explained to them:			
Physical examination by a doctor	8 (82/981)	2 (7/381)	<0.001
X-rays	9 (88/996)	1 (5/379)	<0.001
Ultrasound	9 (39/413)	2 (5/212)	<0.005
Women who wanted more information about the tests they had:			
Physical examination by a doctor	6 (59/964)	2 (7/378)	<0.005
X-rays	9 (68/971)	2 (8/376)	<0.001
Ultrasound	10 (39/401)	3 (6/209)	<0.005

CRC, Cancer Research Campaign; NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Ong *et al.* 1996[111]	**Inclusion criteria**	Literature for women being recalled by UK breast screening assessment centres	**Intervention**	Evaluation of information given in the initial letter/ leaflet and prior to recall for further assessment	**Psychological**	Criteria for evaluating breast screening information material developed by Austoker and Ong 1994[112]
Study design	Cross section	**Exclusion criteria**	NA	***N***	$n = 84$	**Screening attendance**	_
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	$n = 84$	**Control**	NA	**Quality of life**	_
No. of centres	84			***N***	NA		
Length of follow-up	NA						
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign and the NHSBSP						
Conflicts of interest	Not reported						
Notes							
Definition of false-positive	Not reported						
Aim	To evaluate the health education literature for recalled women using criteria developed by Austoker and Ong 1994[112]						
Methodological issues							
Allocation to groups	NA						
Data analysis	Not reported						
Handling missing data	Not reported	**Ethics approval**	NA				
Power calculation	NA						
Subgroup analysis	No						

Demographics					
NA					
Results					
Topics relating to further investigation in the initial written materials inviting women for mammography					
Topic	**Mentioned in any of the written information**				**Mentioned in neither leaflet nor letter nor GP letter**
	In the letter % (*n*) of centres	**In GP letter % (*n*) of centres**	**In the leaflet % (*n*) of centres**	**In both letter and leaflet % (*n*) of centres**	**% (*n*) of centres**
Possibility of recall	46 (39/84)	5 (4/84)	99 (83/84)	45 (38/84)	1 (1/84)
The word 'cancer'[a]	1 (1/84)	4 (3/84)	52 (44/84)	1 (1/84)	49 (41/84)
Particularly worrying information in the recall letter or leaflet					
Topics mentioned	**Mentioned in any of the written information**			**Mentioned in neither recall leaflet nor recall letter**	
	In recall letter % (*n*) of centres	**In recall leaflet % (*n*) of centres**	**In both recall leaflet and letter % (*n*) of centres**	**% (*n*) of centres**	
One or more worrying items:	43 (35/82)	18 (15/82)	7 (6/82)	46 (38/82)	
Word 'cancer'	9 (7/82)	10 (8/82)	1 (1/82)	83 (68/82)	
Words 'treatment', 'something wrong', 'abnormality', or 'abnormal area of the breast'	20 (16/82)	4 (3/82)	1 (1/82)	78 (64/82)	
Word 'hospital'[b]	10 (8/82)	1 (1/82)	0	89 (73/82)	
Words 'not to worry'[c]	22 (18/82)	1 (1/82)	0	77 (63/82)	
Phrase 'nurse counsellor'	5 (4/82)	9 (7/82)	0	87 (71/82)	

Particularly stress-relieving information in the recall letter or leaflet

Topics mentioned	Mentioned in any of the written information			Mentioned in neither recall leaflet nor recall letter
	In recall letter % (*n*) of centres	In recall leaflet % (*n*) of centres	In both recall leaflet and letter % (*n*) of centres	% (*n*) of centres
One or more stress-relieving messages:	68 (56/82)	33 (27/82)	20 (16/82)	17 (14/82)
Most recalled women are found to have normal breasts	28 (23/82)	6 (5/82)	4 (3/82)	30 (25/82)
Recall is part of second stage/routine screening	46 (77/82)	26 (3/82)	11 (9/82)	38 (31/82)
A substantial number of women are recalled	32 (26/82)	11 (9/82)	1 (1/82)	60 (49/82)

CRC, Cancer Research Campaign; NA, not applicable.

a Only when the word 'cancer' was mentioned when referring to further investigation (recall).

b 'Hospital' was only counted when it was mentioned other than in the context of address or directions.

c Similar phrases counted were 'not to be alarmed', 'not to be concerned', 'not to feel anxious', 'no cause for concern'.

Design		Participants		Arms		Outcomes	
Author and year	Austoker and Ong 1994[112]	**Inclusion criteria**	Women invited for routine screening by mammography who were recalled for assessment	**Intervention**	Women invited for routine screening by mammography who were recalled for assessment	**Psychological**	Ad hoc questionnaire including open questions to assess the reassuring or worrying nature of the content of recall letters and leaflets. They were also assessed for coverage of, reason for recall, way to the centre, who could come with them, how to change the appointment, how long it would be, who they would see, what tests would be carried out, when the results would be available and how to get more information
Study design	Cross section	**Exclusion criteria**	Not reported	***N***	*n* = 1493	**Screening attendance**	_
Study centre	CRC Primary Care Education Research Group, University of Oxford	***N***	*n* = 1493	**Control**	NA	**Quality of life**	_
No. of centres	8			***N***	NA		
Length of follow-up	NA						
Setting	NHSBSP clinics						
Funding	Cancer Research Campaign and the NHSBSP						
Conflicts of interest	Not reported						
Notes							
Definition of false-positive	Screened women who underwent further assessment and were found to have nothing wrong						
Aim	To assess the written information needs of women recalled for further assessment						

Methodological issues			
Allocation to groups	NA		
Data analysis	Contingency tables were used for comparison, with a two-sided, $p<0.05$ significance level		
Handling missing data	Not reported	**Ethics approval**	Yes
Power calculation	Not reported		
Subgroup analysis	No		
Demographics			
Not reported			
Results			
How women felt when they received the recall letter			
Reaction	***N* (%) women**	**Sample comments**	
Pleased	30 (2.0)	Very pleased to think I was having a proper check	
Neutral/not distressed	87 (5.9)	I just felt normal	
Somewhat distressed	497 (33.9)	Concerned though not unduly	
		I felt rather apprehensive	
		Nervous, but I think it is a good thing	
		Unpleasantly apprehensive	
Distressed	415 (28.3)	Nervous and very apprehensive	
		Anxious and worried	
		Frightened and worried	
		Worried, afraid	
Very distressed	439 (29.9)	I felt the whole bottom had fallen out of my world	
		I felt sick then faint, then I cried then I kept thinking what I have to do if I have cancer	
		Worried to death	
		Panic stricken, depressed. Convinced I was going to die	
		Completely devastated. Reason abandoned me	
All women	1468 (100)		

Reported need for more information: whether the topic was mentioned or not

Topic	Topic mentioned in letter/leaflet		Topic not mentioned in letter/leaflet		*p*-value
	% (*N*) women wanting more information		% (*N*) women wanting more information		
Why they were recalled	36	(383/1070)	46	(179/388)	<0.005
How to get to the centre	8	(71/854)	26	(75/290)	<0.0001
Who could come with them	5	(44/888)	35	(148/419)	<0.0001
How to change the appointment	2	(33/1436)	–		
How long the appointment would take	8	(17/222)	28	(248/900)	<0.0001
Who they would see	13	(168/1266)	33	(62/186)	<0.0001
What tests would be done	11	(65/606)	35	(298/847)	<0.0001
How to get more information	18	(143/783)	33	(212/633)	<0.0001

Reported need for more information and level of distress

Topic	Distressed/very distressed women		Somewhat/not distressed women		*p*-value
	% (*N*) women wanting more information		% (*N*) women wanting more information		
Why they were recalled	48	(403/834)	26	(157/598)	<0.0001
How to get to the centre	13	(83/659)	13	(64/497)	NS
Who could come with them	13	(102/762)	17	(94/557)	NS
How to change the appointment	2	(18/824)	3	(15/523)	NS
How long the appointment would take	27	(173/640)	20	(93/466)	0.007
Who they would see	18	(146/828)	13	(80/598)	0.030
What tests would be done	27	(224/828)	22	(130/599)	0.022
How to get more information	29	(237/811)	19	(116/616)	<0.0001

Preparing women in advance for possible recall

	Possibility of recall mentioned in the initial screening invitation, % (*N*)	Possibility of recall not mentioned in the initial screening invitation, % (*N*)	*p*-value
Distressed/very distressed women	23 (110/485)	30 (59/197)	<0.05

CRC, Cancer Research Campaign; NA, not applicable.

Design		Participants		Arms		Outcomes	
Author and year	Smith *et al.* 1991[113]	**Inclusion criteria**	Women attending assessment clinic following recall from routine mammography screening	**Intervention**	Three different versions of a recall letter giving increasing amounts of information. Letter two also gave contact details of a BCN	**Psychological**	Ad hoc questionnaire
Study design	Cross section	**Exclusion criteria**	Not reported	***N***	$n = 103$	**Screening attendance**	_
Study centre	Department of Community Health, University of Leicester	***N***	$n = 103$	**Control**	NA	**Quality of life**	_
No. of centres	1			***N***	NA		
Length of follow-up	NA						
Setting	Leicestershire Breast Screening Service						
Funding	Not reported						
Conflicts of interest	Not reported						
Notes							
Definition of false-positive	Not reported						
Aim	To test three different forms of recall letter and to develop and test an audit questionnaire						
Methodological issues							
Allocation to groups	NA						
Data analysis	Fisher's exact test or chi-squared tests were used						
Handling missing data	Not reported	**Ethics approval**	Not reported				
Power calculation	Not reported						
Subgroup analysis	No						

Demographics

Participants were from a predominantly white working class and middle class area

Results

How women felt when they received their invitation letter to return for further tests

Reaction	***N***	**%**
Positive (e.g. 'glad to be in such capable hands')	4	4
Neutral (e.g. 'I wasn't bothered')	10	10
Surprised	11	11
Upset (e.g. 'anxious', 'worried', 'upset')	44	44
Very upset (e.g. 'terrified', extremely anxious')	31	31
Total	100	100

Satisfaction of women with information about why they had to return to clinic and what would happen there

Letter version	**Satisfaction with information on:**	
	Reasons for recall, *N* (%)	**Events at the clinic, *N* (%)**
1	15 (50)	17 (63)
2	25 (71)	24 (74)
3	26 (81)	27 (90)
All versions	66 (68)	68 (76)
Chi-squared	7.243	5.817
p-value	0.027	0.055

Whether recalled women wanted to talk to the BCN

Letter version	Answer	*N*
1. Would telephone the BCN	Yes	25
	No	0
2. Did telephone the BCN	Yes	13
	No	0
3. Would telephone the BCN	Yes	17
	No	1

BCN, breast care nurse; NA, not applicable.

Appendix 4 Ongoing studies

Ongoing trials search: the psychological consequences of false-positive mammograms.

Sources searched for ongoing trials: UKCRN, ControlledTrials.com, ClinicalTrials.gov, WHO ICTRP, www.nhs.uk/Conditions/Clinical-trials/Pages/clinical-trial.aspx, DUETs.

Searches found one UK ongoing study, which is probably outside this scope:

http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=8080
Predicting Risk Of Cancer At Screening (PROCAS).
Breast cancer risk assessment and validation in the National Breast Screening Programme.
Topic.
Cancer (co-adopted by Congenital Disorders).
Portfolio eligibility.
Funded by UK Cancer Research Campaign (CRC), CRC partner.
Current status: open.
Closure date: 1 June 2012.
Not many details – probably not relevant.

Four other non-UK studies were found, which may be of general interest:

http://clinicaltrials.gov/show/NCT00742755
No specific mention of false-positive.
Increasing Adherence to Follow-up of Breast Abnormalities in Low-Income Korean American Women.
This study has been completed.
Study type: interventional.
Study design: allocation, randomised.
End point classification: safety/efficacy study.
Intervention model: parallel assignment.
Masking: single blind (subject).
Primary purpose: screening.
Enrolment: 160.
Study start date: September 2003.
Study completion date: September 2009.
The purpose of this study is to design an intervention to assist Korean American women who have been identified with a *potential breast abnormality* through the Breast Cancer Early Detection Program and who have missed their first follow-up appointment (at risk women). Intervention activities will include reminder telephone calls or home visits by a trained peer navigator to explain the importance of follow-up procedures, emotional support, help with transportation to follow-up appointments, translations, organising care for children or grandchildren during medical appointments, and other assistance to overcome barriers to follow-up identified during the initial phase of the study. The investigators will collect extensive process measures including number and type of intervention activities requested and delivered in order to estimate the feasibility for institutionalising intervention activities. The investigators will conduct chart reviews and a follow-up survey to evaluate the effectiveness of the intervention in increasing adherence to follow-up procedures.
Included as it said 'Potential breast abnormality' – so could result in a false-positive on re-examination.

http://clinicaltrials.gov/ct2/show/NCT00247442
Australian Screening Mammography Decision Aid Trial (ASMDAT).

Not specifically false-positive but interventions might be of interest.

http://clinicaltrials.gov/ct2/show/NCT01267110 – USA study.
Not false-positive – included only for educational to ethnic intervention information.
Engaging Diverse Underserved Communities to Bridge the Mammography Divide.
Estimated enrolment: 242.
Study start date: September 2010.
Estimated study completion date: June 2015.

http://clinicaltrials.gov/show/NCT01261520
Not false-positive – routine mammography – video intervention.
Chinese Women and Mammography Screening.
Enrolment: 671.
Study start date: July 2005.
Primary completion date: August 2009 (final data collection date for primary outcome measure).

Publications:

Wang JH, Mandelblatt JS, Liang W, Yi B, Ma IJ, Schwartz MD. Knowledge, cultural, and attitudinal barriers to mammography screening among nonadherent immigrant Chinese women: ever versus never screened status. *Cancer* 2009;**115**:4828–38.

Sources searched for ongoing trials 26 March 2012: UKCRN, ControlledTrials.com, clinicaltrials.gov, WHO ICTRP.

www.controlled-trials.com/ISRCTN89206644
A clinical trial to investigate the effect of psychological support for women called back for assessment following breast cancer screening: the TLC study is a randomised controlled trial to investigate the effect of psychological support for women called back for assessment following breast cancer screening: the TLC study. Latest information loaded on 21 February 2012.
Primary sponsor: Breast Test Wales (UK).
Date of first enrolment: 1 July 2007.
Target sample size: 300.
Recruitment status: completed/not recruiting.
URL: http://isrctn.org/ISRCTN89206644.
Inclusion criteria: (1) females, aged 50–64; (2) those attending a Breast Test Wales Centre for a recall visit following initial breast screening; (3) participant should be willing to give verbal and written consent for the study; and (4) participant should be willing to complete a questionnaire prior to assessment at baseline, within 1 month, 6 months and 12 months post assessment.
Exclusion criteria: (1) those who are recalled for technical reasons (technical recall); (2) women who have had a previous recall within the last 3 years; (3) women who have any hearing, visual or learning impairment which would not allow them to complete the questionnaires or listen to the support package; (4) women who themselves have identified breast problems (clinical override); and (5) women who cannot answer questionnaires in English or Welsh.

Outcomes:

Primary: score on the negative subscale of the PCQ. All primary and secondary outcomes will be assessed at baseline, 6 weeks, 6 months and 1 year.

Secondary: (1) SF-36 Health Survey; (2) HADS; (3) Euroquol EQ-5D; and (4) Short Explanatory Model Interview for patient experiences. All primary and secondary outcomes will be assessed at baseline, 6 weeks, 6 months and 1 year.

Intervention: some women invited for breast screening are then asked to attend for further tests. This study looks at a relaxation and self-help package known as 'Travel Lightly Companion' (TLC) to see if it reduces any distress linked to recall. The TLC pack consists of guided self-help presented as a CD of relaxation music with relaxation exercises including breathing and guided imagery exercises. Women agreeing to take part will get either the TLC package or care as usual. Participants fill out some questionnaires at the start, 6 weeks, 6 months and 1 year later.

Conference searches:

7th European Breast Cancer Conference (EBCC-7).
Barcelona, Spain, 24–27 March 2010 – nothing relevant.

First British Breast Cancer Research Conference.
East Midlands Conference Centre, Nottingham, UK.
15–17 September 2010 (planned to be held biennially).
www.bbcrc.org.uk/bbx/contact.asp
No abstracts online only programme.

15th Annual Mammography Update for Physicists (2011).
www.mtmi.net/seminars/mam_phys_update.php

American Society of Clinical Oncology (ASCO) Breast Cancer Symposium 2010 – nothing relevant.

11th St. Gallen International Breast Cancer Conference – nothing relevant.

Society of Radiographers.
www.sor.org/

Radiological Society of North America (RSNA) 2010 – nothing relevant for conference.
www.rsna.org/media/pressreleases/pr_target.cfm?ID=475
www.rsna.org/Publications/rsnanews/oct05/digitalmammography.cfm

Appendix 5 Papers excluded at full review

	Citation	Reason excluded
1.	Absetz P, Aro AR, Sutton SR. Experience with breast cancer, pre-screening perceived susceptibility and the psychological impact of screening. *Psychooncology* 2003;**12**:305–18.	Setting only
2.	Andersen SB, Vejborg I, von Euler-Chelpin M. Participation behaviour following a false positive test in the Copenhagen mammography screening programme. *Acta Oncol* 2008;**47**:550–5.	Setting only
3.	Andrykowski MA, Carpenter JS, Studts JL, Cordova MJ, Cunningham LLC, Beacham A, *et al*. Psychological impact of benign breast biopsy: a longitudinal, comparative study. *Health Psychol* 2002;**21**:485–94.	Setting only
4.	Aro AR, Pilvikki Absetz S, van Elderen TM, van der Ploeg E, van der Kamp LJ. False-positive findings in mammography screening induces short-term distress – breast cancer-specific concern prevails longer. *Eur J Cancer* 2000;**36**:1089–97.	Setting only
5.	Austoker J. Screening and self examination for breast cancer. *BMJ* 1994;**309**:168–74.	Design
6.	Baines CJ, To T, Wall C. Women's attitudes to screening after participation in the National Breast Screening Study. A questionnaire survey. *Cancer* 1990;**65**:1663–9.	Setting only
7.	Barton MB, Moore S, Morley DS, Allen JD, Kleinman KP, Emmons KE, *et al*. Decreasing anxiety after false-positive mammograms: a controlled trial. *J Gen Intern Med* 2002;**17**:141.	Setting only
8.	Barton MB, Morley DS, Moore S, Allen JD, Kleinman KP, Emmons KM, *et al*. Decreasing women's anxieties after abnormal mammograms: a controlled trial. *J Natl Cancer Inst* 2004;**96**:529–38.	Setting only
9.	Barton MB, Moore S, Polk S, Shtatland E, Elmore JG, Fletcher SW. Increased patient concern after false-positive mammograms – clinician documentation and subsequent ambulatory visits. *J Gen Intern Med* 2001;**16**:150–6.	Setting only
10.	Baskinsmith J, Miaskowski C, Dibble SL, Weekes D, Nielsen BB. Perceptions of the mammography experience. *Cancer Nurs* 1995;**18**:47–52.	Setting, population
11.	Bishop C, Fisher EB, Heckman BD, Merbaum M, Monsees B, Ristvedt S. Coping and anxiety in women recalled for additional diagnostic procedures following an abnormal screening mammogram. *Health Psychol* 2004;**23**:42–8.	Duplicate
12.	Boekema AG, Dornseiffen G, Mulder HJ, de Vos RA, Kluft-de Haas BA. Initial results with breast screening in the Enschede area. II. Efficiency. *Ned Tijdschr Geneeskd* 1992;**136**:1761–4.	Language, population, setting
13.	Brewer NT, Salz T, Lillie SE. Research digest: false-positive mammogram results may affect feelings, actions. *Patient Care* 2007;**41**:2.	Design
14.	Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. *J Med Screen* 2004;**11**:39–44.	Design
15.	Brown ML, Chiu Y, Currence B, Moore A, Earp J, Pisano ED. A telephone survey to determine the psychosocial effects of a false positive mammogram leading to a breast biopsy vs short-term follow-up mammography. *Radiology* 2002;**225**:369.	Setting, design
16.	Burman ML, Taplin SH, Herta DF, Elmore JG. Effect of false-positive mammograms on interval breast cancer screening in a health maintenance organization. *Ann Int Med* 1999;**131**:1–6.	Setting only
17.	Byram SJ. *Breast cancer and mammogram screening: mental models and quantitative assessments of beliefs*. ProQuest Information & Learning; 1999.	Unobtainable thesis
18.	Calnan MW, Moss S, Chamberlain J. Explaining attendance at a breast screening clinic. *Patient Educ Couns* 1985;**7**:87–96.	Population, length of follow up

Citation		Reason excluded
19.	del Carmen C, Rodriguez A, Cruzado JA, Gonzalez A, Gonzalez MJ, Roa A. Anxiety, depression and quality of life in the course of mammography screening among women of the Madrid Region. *Psicooncologia* 2008;**5**:129–46.	Language, population, setting
20.	Chiarelli AM, Moravan V, Halapy E, Majpruz V, Mai V, Tatla RK. False-positive result and reattendance in the Ontario Breast Screening Program. *J Med Screen* 2003;**10**:129–33.	Setting only
21.	Champion V. Development of a benefits and barriers scale for mammography utilization. *Cancer Nurs* 1995;**18**:53–9.	Population, outcomes, setting
22.	Champion V, Huster G. Effect of interventions on stage of mammography adoption. *J Behav Med* 1995;**18**:169–87.	Intervention, population, outcomes, setting
23.	Clutton S, Pakenham KI, Buckley B. Predictors of emotional well-being following a 'false-positive' breast cancer screening result. *Psychol Health* 1999;**14**:263–75.	Setting only
24.	Cockburn J, Staples M, Hurley SF, De Luise T. Psychological consequences of screening mammography. *J Med Screen* 1994;**1**:7–12.	Setting only
25.	Currence BV, Pisano ED, Earp JA, Moore A, Chiu YF, Brown ME, *et al*. Does biopsy, aspiration or six-month follow-up of a false-positive mammogram reduce future screening or have large psychosocial effects? *Acad Radiol* 2003;**10**:1257.	Setting only
26.	DeFrank JT, Brewer N. A model of the influence of false-positive mammography screening results on subsequent screening. *Health Psychol Rev* 2010;**4**:112–17.	Outcomes, setting, design
27.	Ekeberg O, Skjauff H, Karesen R. Screening for breast cancer is associated with a low degree of psychological distress. *Breast* 2001;**10**:20–4.	Setting only
28.	Eardley A, Elkind A. Breast screening among women aged 65 and over – what do they think of it? *J Public Health Med* 1991;**13**:172–7.	Population
29.	Edwards Adrian GK, Evans R, Dundon J, Haigh S, Hood K, Elwyn Glyn J. *Personalised risk communication for informed decision making about taking screening tests*. In Cochrane Database of Systematic Reviews. Issue 4. Chichester: John Wiley & Sons, Ltd; 2006.	Intervention, population, comparator, outcomes
30.	Elkind A, Eardley A. Consumer satisfaction with breast screening – a pilot study. *J Public Health Med* 1990;**12**:15–18.	Population
31.	Fallowfield LJ, Rodway A, Baum M. What are the psychological factors influencing attendance, nonattendance and re-attendance at a breast screening center. *J R Soc Med* 1990;**83**:547–51.	Population
32.	Ganott MA, Sumkin JH, King JL, Klym AH, Catullo VJ, Cohen CS, *et al*. Screening mammography: do women prefer a higher recall rate given the possibility of earlier detection of cancer? *Radiology* 2006;**238**:793–800.	Population, outcomes
33.	Gibson CJ, Weiss J, Goodrich M, Onega T. False-positive mammography and depressed mood in a screening population: findings from the New Hampshire Mammography Network. *J Pub Health* 2009;**31**:554–60.	Setting only
34.	Gilbert FJ, Cordiner CM, Affleck IR, Hood DB, Mathieson D, Walker LG. Breast screening: the psychological sequelae of false-positive recall in women with and without a family history of breast cancer. *Eur J Cancer* 1998;**34**:2010–14.	Population
35.	Godby CJ, Peticolas A, Carter JM, Ghebrehewet S, Woloshin S, Schwartz L, *et al*. Women's attitudes to false positive mammography results. 2000;**321**:1409–11.	Design
36.	Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *The Cochrane Collaboration*. 2006;**4**.	Unobtainable
37.	Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* 2009;**4**:	Outcomes
38.	Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* 2011;**1**.	Outcomes

Citation		Reason excluded
39.	Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. *Br J Cancer* 1990;**62**:1018–22.	Setting only
40.	Gram IT, Slenker SE. Cancer anxiety and attitudes toward mammography among screening attenders, nonattenders, and women never invited. *Am J Public Health* 1992;**82**:249–51.	Setting, length of follow-up
41.	Haas J, Kaplan C, McMillan A, Esserman LJ. Does timely assessment affect the anxiety associated with an abnormal mammogram result? *J Womens Health Gend Based Med* 2001;**10**:599–605.	Population, outcomes, setting
42.	Heckman BD, Fisher EB, Monsees B, Merbaum M, Ristvedt S, Bishop C. Coping and anxiety in women recalled for additional diagnostic procedures following an abnormal screening mammogram. *Health Psychol* 2004;**23**:42–8.	Setting only
43.	Heckman BD. *Psychosocial factors that promote adjustment among women recalled for a diagnostic assessment following a screening mammogram*. ProQuest Information & Learning; 2001.	Unobtainable thesis
44.	Hislop TG, Harris SR, Jackson J, Thorne SE, Rousseau EJ, Coldman AJ, *et al*. Satisfaction and anxiety for women during investigation of an abnormal screening mammogram. *Breast Cancer Res Treat* 2002;**76**:245–54.	Outcomes, population, setting
45.	Hoekstra SA, Gentry D. Benefits of same day mammogram results. *Am J Clin Oncol* 2009;**32**:552–3.	Setting, design, outcomes, population
46.	Hofvind S, Thoresen S, Tretli S. The cumulative risk of a false-positive recall in the Norwegian Breast Cancer Screening Program. *Cancer* 2004;**101**:1501–7.	Setting only
47.	Jatoi I, Zhu K, Shah M, Lawrence W. Psychological distress in U.S. women who have experienced false-positive mammograms. *Breast Cancer Res Treat* 2006;**100**:191–200.	Setting only
48.	Kahn BE, Luce MF. Repeated-adherence protection model: 'I'm OK, and it's a hassle'. *J Public Policy Mark* 2006;**25**:79–89.	Setting design
49.	Lampic C, Thurfjell E, Bergh J, Sjoden PO. Short- and long-term anxiety and depression in women recalled after breast cancer screening. *Eur J Cancer* 2001;**37**:463–9.	Setting only
50.	Lampic C, Thurfjell E, Sjoden PO. The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. *Eur J Cancer* 2003;**39**:1730–7.	Setting only
51.	Lerman C, Ross E, Boyce A, Gorchov PM, McLaughlin R, Rimer B, *et al*. The impact of mailing psychoeducational materials to women with abnormal mammograms. *Am J Public Health* 1992;**82**:729–30.	Setting only
52.	Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. *Ann Intern Med* 1991;**114**:657–61.	Setting only
53.	Lidbrink E, Levi L, Pettersson I, Rosendahl I, Rutqvist LE, Delatorre B, *et al*. Single-view screening mammography – psychological, endocrine and immunological effects of recalling for a complete 3-view examination. *Eur J Cancer* 1995;**31A**:932–3.	Setting only
54.	Lipkus IM, Halabi S, Strigo TS, Rimer BK. The impact of abnormal mammograms on psychosocial outcomes and subsequent screening. *Psychooncology* 2000;**9**:402–10.	Setting only
55.	Lowe JB, Balanda KP, Del Mar C, Hawes E. Psychologic distress in women with abnormal findings in mass mammography screening. *Cancer* 1999;**85**:1114–18.	Population, setting
56.	Medical Advisory Secretariat. *Screening mammography for women aged 40 to 49 years at average risk for breast cancer: an evidence-based analysis*. 2007. URL: www.health.gov.on.ca/english/providers/program/mas/mas_mn.html	Outcomes, setting, population
57.	Miller EG, Luce MF, Kahn BE, Conant EF. Understanding emotional reactions for negative services: the impact of efficacy beliefs and stage in process. *J Serv Res* 2009;**12**:87–99.	Population, setting

Citation		Reason excluded
58.	Olsson P, Armelius K, Nordahl G, Lenner P, Westman G.Women with false positive screening mammograms: how do they cope?. *J Med Screen* 1999;**6**:89–93.	Setting only
59.	Padgett DK, Yedidia MJ, Kerner J, Mandelblatt J. The emotional consequences of false positive mammography: African-American women's reactions in their own words. *Women Health* 2001;**33**:1–14.	Setting only
60.	Pinckney RG, Geller BM, Burman M, Littenberg B. Effect of false-positive mammograms on return for subsequent screening mammography. *Am J Med* 2003;**114**:120–5.	Setting only
61.	Pisano ED, Earp JA, Gallant TL. Screening mammography behavior after a false positive mammogram. *Cancer Detect Prev* 1998;**22**:161–7.	Setting only
62.	Pisano ED, Earp J, Schell M, Vokaty K, Denham A. Screening behavior of women after a false-positive mammogram. *Radiology* 1998;**208**:245–9.	Setting only
63.	Rimer BK, Keintz MK, Kessler HB, Engstrom PF, Rosan JR. Why women resist screening mammography – patient related barriers. *Radiology* 1989;**172**:243–6.	Population, setting
64.	Scaf-Klomp W, Sanderman R, Van De Wiel HBM, Otter R, Van Den Heuvel WJA. Distressed or relieved? Psychological effects of participation in the population screening for breast cancer milder than expected. *Ned Tijdschr Geneeskd* 1999;**143**:459–64.	Language, setting
65.	Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. *J Epidemiol Community* 1997;**51**:705–10.	Setting only
66.	Stewart KA, Neumann PJ, Fletcher SW, Barton MB. The effect of immediate reading of screening mammograms on medical care utilization and costs after false-positive mammograms. *Health Serv Res* 2007;**42**:1464–82.	Setting only
67.	Stewart KA, Neumann PN, Fletcher SW, Barton MB. Impact on utilization and cost of an intervention to decrease anxiety after false-positive mammograms. *J Gen Intern Med* 2005;**20**:212.	Outcomes, setting, design
68.	Tobias IS, Baum M. False positive findings of mammography will have psychological consequences. *BMJ* 1996;**312**:1227.	Outcomes, setting, design
69.	van Wijngaarden DA. Distressed or relieved? Psychological sequelae of breast cancer screening in Netherlands. *Ned Tijdschr Geneeskd* 1999;**143**:873–4.	Language, setting
70.	Wardle J, Pope R. The psychological costs of screening for cancer. *J Psychosom Res* 1992;**36**:609–24.	Design
71.	Zapka JG, Puleo E, Taplin SH, Goins KV, Ulcickas Yood M, *et al*. Processes of care in cervical and breast cancer screening and follow-up – the importance of communication. *Prev Med* 2004;**39**: 81–90.	Setting only

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library

Published by the NIHR Journals Library