Psychological consequences of false-positive screening mammograms in the UK

Mary Bond,[1] Toby Pavey,[1] Karen Welch,[2] Chris Cooper,[1] Ruth Garside,[3] Sarah Dean,[1] Christopher J Hyde[1]

10.1136/eb-2012-100608

[1]PenTAG, University of Exeter, Peninsula Medical School, Exeter, Devon, UK
[2]Karen Welch Information Consultancy, Fareham, UK
[3]European Centre for Environment and Human Health, University of Exeter, Truro, UK

Correspondence to:
Dr Mary Bond
PenTAG, Peninsula Medical School, University of Exeter, Veysey Building, Salmon Pool Lane, Exeter EX2 4SG, Devon, UK;
mary.bond@pms.ac.uk

Abstract

Objectives To identify the psychological effects of false-positive screening mammograms in the UK.

Methods Systematic review of all controlled studies and qualitative studies of women with a false-positive screening mammogram. The control group participants had normal mammograms. All psychological outcomes including returning for routine screening were permitted. All studies had a narrative synthesis.

Results The searches returned seven includable studies (7/4423). Heterogeneity was such that meta-analysis was not possible. Studies using disease-specific measures found that, compared to normal results, there could be enduring psychological distress that lasted up to 3 years; the level of distress was related to the degree of invasiveness of the assessment. At 3 years the relative risks were, further mammography, 1.28 (95% CI 0.82 to 2.00), fine needle aspiration 1.80 (95% CI 1.17 to 2.77), biopsy 2.07 (95% CI 1.22 to 3.52) and early recall 1.82 (95% CI 1.22 to 2.72). Studies that used generic measures of anxiety and depression found no such impact up to 3 months after screening. Evidence suggests that women with false-positive mammograms have an increased likelihood of failing to reattend for routine screening, relative risk 0.97 (95% CI 0.96 to 0.98) compared with women with normal mammograms.

Conclusions Having a false-positive screening mammogram can cause breast cancer-specific distress for up to 3 years. The degree of distress is related to the invasiveness of the assessment. Women with false-positive mammograms are less likely to return for routine assessment than those with normal ones.

Introduction

The benefits and harms arising from mammography screening are a matter of national debate in the UK.[1 2] The number of lives saved, amount of over diagnosis and degree of distress caused by 'false alarms' are hotly contested.[3–7] This debate has led to a review of UK breast cancer screening services currently being undertaken by Professor Sir Michael Richards.

The negative psychological impact of false-positive screening results has been documented in the fields of prenatal and cervical cancer screening.[8 9] Their impact on the medium to long-term psychological well-being and behaviour of women who receive false-positive results from routine mammography has been less well researched and synthesised.

Systematic reviews and meta-analyses in Europe and America have found conflicting evidence about the psychological impact from receiving a false-positive mammogram and also on future attendance at routine screening.[7 10–14] Most studies showed a negative impact from receiving a false-positive mammogram on measures of, well-being, depression and anxiety compared to women with normal screening results. The exception to this was the meta-analysis of generic psychological measures by Salz *et al*[10] which showed that only anxiety 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.[13]

The results for the impact of receiving a false-positive mammogram on returning for the next routine screening mammogram give a more complex picture. Armstrong *et al*[15] 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*[14] 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.[7]

These systematic reviews had a wider geographical scope than the UK and included studies of the short-term impact of a false-positive mammogram (less than 1 month from assessment). Many of the included studies were based on programmes with a different approach and periodicity to that of the UK where mammography screening is national, free, opt-out and uses double screening of mammograms. Additionally the UK service runs on a 3 years cycle of invitations to women aged 47–73. In particular, the US system differs as it comes from a mixture of public and private providers and is insurance-based, opt-in, uses single screening, which produces a higher number of false-positives and has until recently recommended screening annually from the age of 40. Most European countries offer screening for the 50–69 age group but generally every 2 years, and are provided by a mixture of public and private organisations, which may or may not be publically financed. Therefore, the differences in the provision of screening services indicate that the psychological effects of this experience and the impact on returning for screening

remains unclear in the particular health service and cultural context of the UK.

To address this knowledge gap we conducted a systematic review of studies in the UK population of the medium to long-term psychological consequences of experiencing a false-positive screening mammogram and whether these affect future attendance at mammography screening. In addition qualitative studies, an important underpinning to the understanding of psychological consequences, did not appear to have been searched for in previous reviews, so we particularly targeted these in our study.

The definition of a false-positive mammogram used in this study is that given by the WHO: 'an abnormal mammogram (one requiring further assessment) in a woman ultimately found to have no evidence of cancer'[16] and the American College of Radiology categories of mammographies considered as abnormal are detailed in http://www.imaginis.com/mammography/mammogram-interpretation-categories-and-the-acr-bi-rads.

Methods

This systematic review was carried out following the principles published by the NHS Centre for Reviews and Dissemination.[17] The study protocol can be found in appendix 1.

Inclusion criteria

Studies were included if based in the UK, populated by women whose experience met the above definition of a false-positive screening mammogram, the comparator group were those with a normal screening mammogram, the outcomes were psychological, behavioural or those from qualitative studies and follow-up was at least 1 month from the 'all clear'. All controlled studies and qualitative designs were included. Case studies were excluded.

Search strategy

The search strategy consisted of searching of electronic bibliographic databases, internet searches, scrutiny of references of included studies and contacting experts in the field.

The following electronic databases were searched in December 2010: Medline, Medline in Process and other Non-Indexed Citations, EMBASE, HMIC, Cochrane Central, Cochrane CDSR, CRD Dare, CRD HTA, Cochrane Methodology, Web of Science, Psychinfo, Cinahl, Sociological Abstracts, the International Bibliography of the Social Sciences and Zetoc. Ongoing trials were searched for at: UKCRN, Controlled Trials.com, Clinical Trials.gov, ICTRP (WHO International Clinical Trials Registry Platform), 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 more qualitative and grey literature were run in January 2011 on the following databases: Medline in Process and other Non-Indexed Citations, Embase Classic and Embase, British Nursing Index and Archive, Social Policy and Practice, Cinahl plus, Cochrane Library, HMIC, PsycINFO, Assia, Sociological Abstracts, Web of Science, CRD and IBSS. All searches were run from inception to the then present date. Bibliographies of included studies were searched for further relevant studies. An update search was carried out on 26 November 2011. The MEDLINE search strategy is available in appendix 2.

Papers were selected for review from the titles and abstracts generated by the search strategy. This was done independently by two reviewers (MB and TP); discrepancies were resolved by discussion. Retrieved papers were again reviewed and selected against the inclusion criteria by the same independent process. Data were extracted from included studies by one reviewer using standardised data extraction forms and checked by another reviewer. Attempts were made to contact authors to provide missing information. Data were gathered on the design, participants, methods, outcomes, baseline characteristics and results of the studies.

Quality assessment

Studies were assessed for internal validity according to criteria suggested by the updated NHS CRD Report No. 4, according to study type.[17 18] Randomised controlled trials (RCTs) were appraised with the CONSORT statement[19] and observational studies with STROBE guidelines.[20] External validity was assessed according to the applicability of findings to a relevant patient group and service setting.

Analysis

Analysis was carried out using StatSEv12 software. The principle summary measures were relative risks with 95% CI. All study designs had a narrative synthesis. Observational studies were considered for possible meta-analysis. Overall, they had considerable amounts of missing information so that it was difficult to judge heterogeneity. Therefore, a meta-analysis was not carried out.

Results

Search results

Our searches retrieved 4423 titles and abstracts after deduplication. When screening was complete we found seven primary studies (nine papers) that met the inclusion criteria. Four of the studies were prospective cohorts[21–25] four were retrospective cohorts,[26–28] and one was an RCT of an intervention to improve reattendance.[29]

Studies covered two domains: (a) three studies looked at the psychological impact of false-positive mammograms in the normal risk population;[21–25] and (b) six looked at the impact of this experience on returning for routine mammogram screening;[21 22 26–29] (some studies looked at both domains). No studies were found that were either about or that had subgroups of women from different ethnic, socioeconomic or other groups within the general screening population. No published qualitative studies were found. A flow chart of the selection process is in appendix 3.

Quality and characteristics

The quality of the research was variable; the RCT was poor quality with few methodological details given,

although some of the observational studies, notably those from the Oxford Primary Care Education Research Group, were reasonably well reported. However, the majority had a number of weaknesses, including a failure to consider the possible effects of bias and confounding on the results and a failure to report participants' demographic and other characteristics. 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. There were insufficient studies in each domain to produce a meaningful assessment of publication bias with a funnel plot. Table 1 provides summary characteristics of the included studies by their outcome domain. Appendix 4 shows a summary assessment of quality indicators.

Study results

Psychological impact

The studies of the psychological impact of false-positive mammograms gave conflicting results. When disease specific measures were used that is, the Psychological Consequences Questionnaire[30] an enduring negative impact was found for those with false-positive results compared to those with normal mammograms 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 (relative risk (RR) 95% CI 2.07 (1.22 to 3.52)) than women who had fine needle aspiration (RR 95% CI 1.80 (1.17 to 2.77)), and non significantly; further mammography (RR 95% CI 1.28 (0.82 to 2.00)). Additionally women placed on early recall were also at a greater relative risk of distress (RR 95% CI 1.82 (1.22 to 2.72)). The greatest relative risk of distress was felt at 5 months after assessment and was significant for all assessment procedures (figure 1).

Conversely, when generic measures of general anxiety and depression were used, the Hospital Anxiety and Depression Scale[31] (http://www.surreyhealth.nhs.uk/dcp/Documents/D1.3d2.pdf) and the General Health Questionnaire-28[32] no significant differences were found between the two groups at 6 weeks after assessment and 3 months after screening.[25 24] (figure 2).

Reattendance

The evidence for the impact of having a false-positive mammogram on returning for the next screening round is again conflicting. The forest plot below (figure 3) compares the relative risks of the reattendance studies. The evidence comes 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, is that women with false-positive mammograms are less likely to return for their next round of screening than women with normal mammograms, although the effect is small. The largest study with this finding (N=140 387) had a relative risk of returning of 0.97 (95% CI 0.96 to 0.98).[26] Brett and Austoker[22] came to the same conclusion, 0.92 (95% CI 0.86 to 0.98). Two studies with a combined population of N=7231 found that there was no such association but had wide 95% CIs consistent with both increased and decreased likelihood of return.[27 28]

Evidence from a poor-quality RCT by Meldrum *et al*[29] (N=3083) 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 (95% CI) 1.10 (1.00 to 1.21)).

Discussion

The benefits and harms accruing from breast cancer screening are a matter of current UK concern.

Included studies, comparing psychological impact of false-positive mammograms to normal ones, gave conflicting results. When disease specific measures were used an enduring negative impact was found that lasted until 35 months with the degree of distress related to the invasiveness of the assessment. Conversely, when measures of general anxiety and depression were used no significant differences were found between the two groups. However, this could be explicable if we speculate that false-positive mammograms may lead to breast cancer-specific psychological distress, enduring for up to 3 years, but that it is unlikely that general anxiety or depression will occur.

Concerning reattendance, the weight of evidence is that women with false-positive mammograms are less likely to return for subsequent rounds of screening than women with normal mammograms.

No systematic reviews were found that are directly comparable to ours as they all include non-UK studies, may have populations younger than ours, measure outcomes at less than 1 month and have screening programmes based on opt-in or insurance payments.[7 10 11 13–15] Nevertheless, our results agree with theirs that there can be negative psychological consequences from having a false-positive mammogram and that reattendance can fall. Greater clarity has emerged by removing the effect of variation in the specific nature of different national programmes.

Limitations

The robustness of the findings of this systematic review are limited by the reliability of the poorly reported observational studies. The degree of heterogeneity between these studies is subsequently unknown. This meant that we were unable to pool the data (without potentially reporting spurious relationships) and thus provide an overall estimate of distress and reattendance. Additionally, the evidence in this systematic review is at least 10 years old and may also have been influenced by publication bias.

Our decision to restrict included studies to those in the UK may also be seen as a limitation. However we felt that this was a reasonable approach given the variability in results worldwide, the most obvious explanation for which was likely to be variation in programme. Detailed investigation of qualitative research was also felt to be more appropriately conducted at national level. Also despite the UK-specific nature of this review we believe that generalisable messages remain, as indicated below.

Table 1 Summary characteristics of included studies by outcome 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 (Cancer Research Campaign)	Prospective cohort multicentre *Psychological impact*	505	Women invited for routine screening by mammogram, already participating in the study at 5 months	Routine screening by mammogram with a false-positive result N=375	Routine screening by mammogram 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	Over 65 years, symptomatic referral, in another study, developed cancer	
Brett *et al*, 1998 (Cancer Research Campaign)	Prospective cohort multicentre *Psychological impact*	284	Women invited for routine screening by mammogram, already participating in the study at 1 month	Routine screening by mammogram with a false-positive result N=163	Routine screening by mammogram with a normal result N=52	PCQ, intention to reattend, ad hoc questionnaire	5 months after assessment	As above	69 (24%) women chose not to return the questionnaire
Ong *et al*, 1997a (Cancer Research Campaign, NHSBSP)	Prospective cohort multicentre *Psychological impact*	877	Women invited for routine screening by mammogram 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	1 month after assessment	Not reported	This study was primarily about the effects of early recall on women who had been recalled after their mammogram
Bull and Campbell, 1991 (Yorkshire Regional Health Authority)	Prospective cohort *Psychological impact*	750	Women invited for routine screening by mammogram recalled for assessment	Routine screening by mammogram with a false-positive result N=308	Routine screening by mammogram 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
Ellman *et al*, 1989 (DHSS Research Management Division)	Prospective cohort *Psychological impact*	752	Women invited for routine mammogram screening and those recalled for further assessment and those with symptoms being further investigated	Routine screening by mammogram with a false-positive result N=271	Routine screening by mammogram 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
Impact on reattendance									
McCann *et al*, 2002 (NHS Executive Eastern Region)	Retrospective cohort *Reattendance and interval cancer*	140387	Women 49–63 years invited for routine breast screening by mammography	Routine screening by mammogram with a false-positive result N=4 792	Routine screening by mammogram 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	

Continued

Table 1 Continued

Study/author year (funding)	Design	N	Participants	Intervention group	Control group	Outcomes	Length of follow-up	Exclusion criteria	Notes
O'Sullivan *et al*, 2001 (Cancer Research Campaign)	Retrospective cohort *Reattendance*	5649	Women invited for mammography screening for the second or more time	Routine screening by mammogram with a false-positive result N=248	Routine screening by mammogram 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 (Cancer Research Campaign)	As above in psychological impact								
Brett *et al*, 1998 (Cancer Research Campaign)	As above in psychological impact								
Meldrum, 1994 (Scottish Office Home and Health Department)	RCT-nested telephone interview study	3083	All women invited for second round routine mammography screening (50–65 years)	Tailored invitation with a false-positive result N=115 and with 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
Orton, 1991 (funding not reported)	Retrospective cohort *Reattendance*	1582	Women, aged 45–64, invited to attend for second round screening by mammography	Routine screening by mammogram with a false-positive result N=50	Routine screening by mammogram 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

FNA, fine needle aspiration; HADS, Hospital Anxiety and Depression Scale; NA, not applicable; PCQ, Psychological Consequences Questionnaire; RCT, randomized controlled trial.

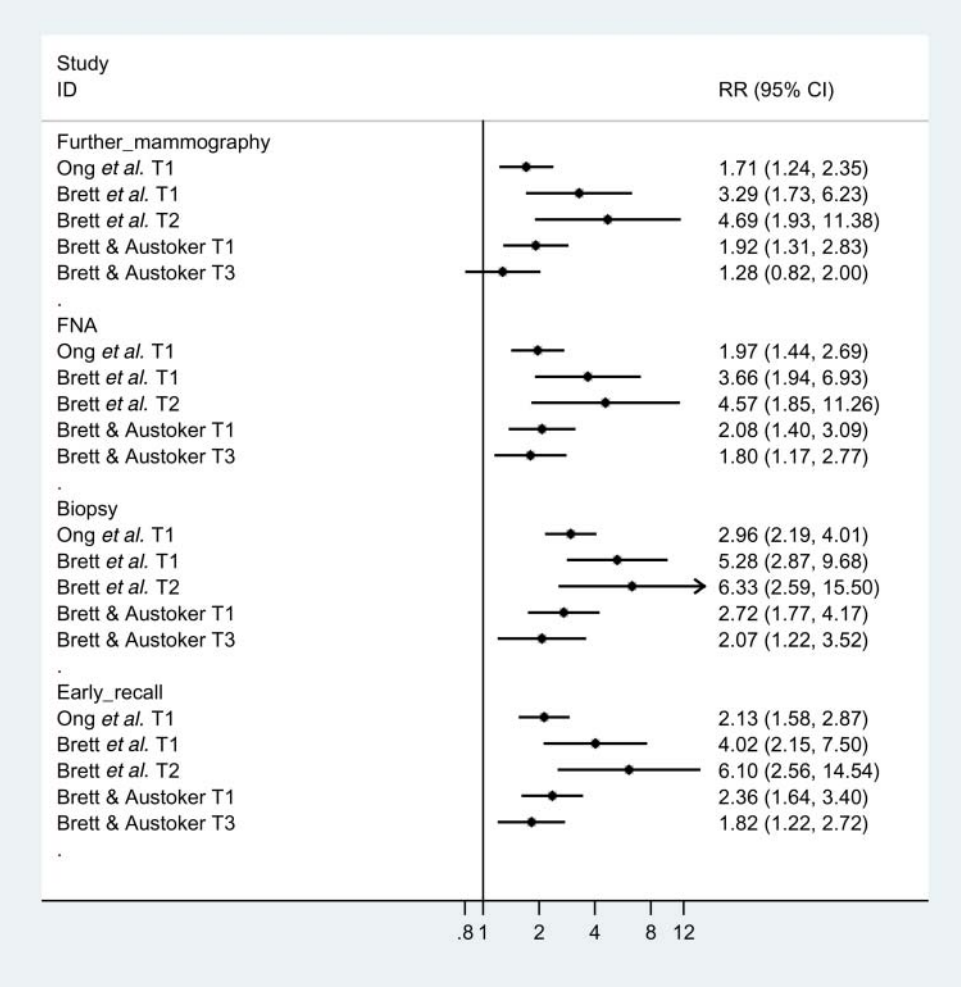


Figure 1 Forest plot of the relative risks of negative psychological consequences from having a false-positive mammogram compared to a normal one by type of false-positive assessment, at T1 (1 month after assessment), T2 (5 months after assessment) and T3 (35 months after assessment), measured with the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

Implications for policy and future research

Policy makers, particularly in the UK should consider the impact of false-positive mammograms when planning breast screening services. Measures need to be taken to reduce the distress caused; however, the evidence base for such measures is lacking. Internationally the need to take account of the context of the evaluations is illustrated and absence of supporting qualitative research to understand the underlying nature of psychological consequences in the UK is likely to have parallels in other countries.

Therefore, further research is needed to increase and update our understanding of the harms of breast cancer screening. In particular a qualitative interview study would further our appreciation of what this experience means to women, and principally shed light on how the subtleties of difference in response to assessment procedures relate to anxiety and probability of reattendance. Consequently, the authors are currently conducting such an interview study. Well-designed observational studies are also needed, that use disease specific and generic outcome measures in order to determine the degree and kind of negative psychological outcomes. Studies should include women from different ethnic and socio-economic groups and routinely collect demographic information so that future systematic reviews may be able to judge whether the pooling of data is legitimate.

Conclusions

We conclude that the experience of having a false-positive screening mammogram can cause breast cancer-specific psychological distress that may endure for 3 years. However, it is less likely that there will be pathological general anxiety and depression. The distress caused by a false-positive mammogram may be

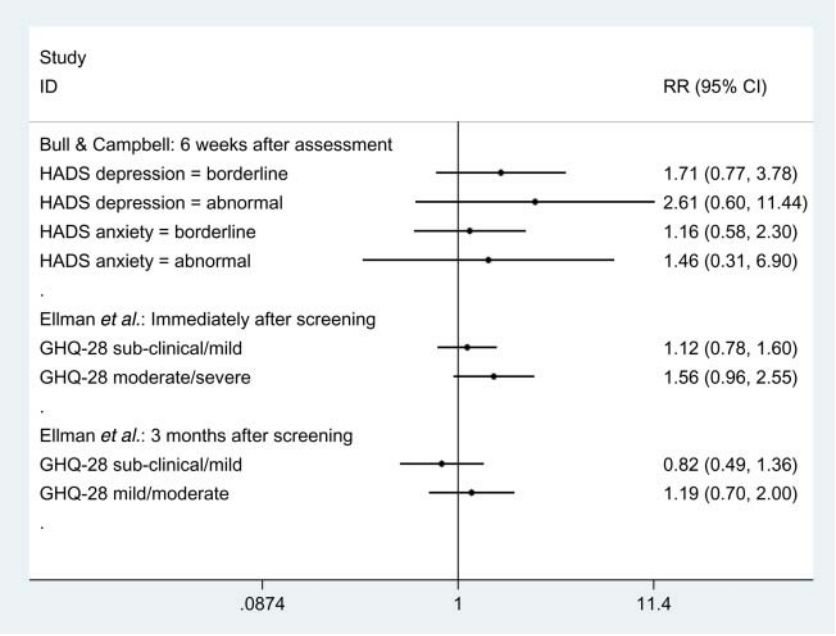


Figure 2 Relative risk of suffering clinically measurable levels of general anxiety and depression following a false-positive mammogram compared to a normal mammogram, measured by Bull and Campbell (HADS) and Ellman *et al* (GHQ-28).

sufficient to deter some women from attending their next breast cancer screening appointment.

▶ Additional data is published online only. To view this file please visit the journal online (http://ebm.bmj.com).

Funding This project was funded by the NIHR Health Technology Assessment programme (project number 09/145/01) and will be published in full in the *Health Technology Assessment* journal, Vol. [x], No. [x]. See the HTA programme website for further project information. SD was funded by the NIHR through the Peninsula Collaboration for Leadership in Applied Health Research and Care.

Disclaimer The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

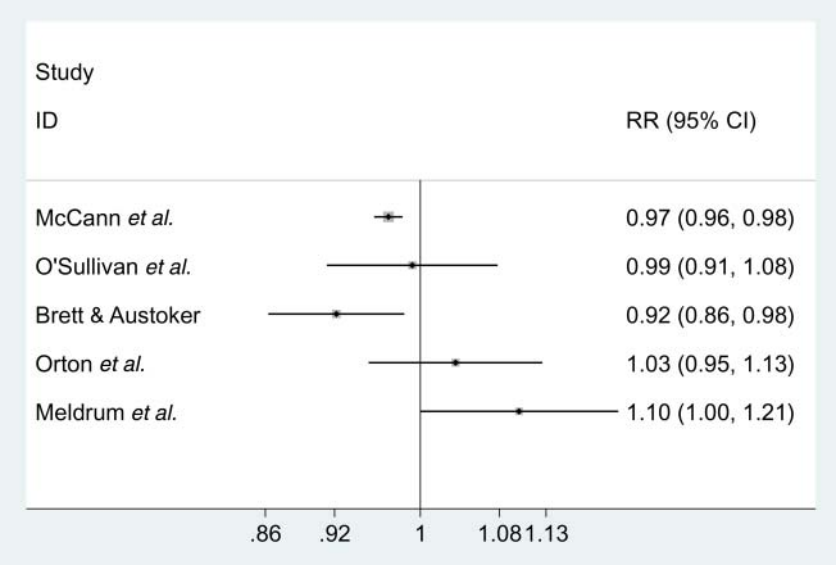


Figure 3 Forest plot of the likelihood of failing to reattend the next round of mammography screening following a false-positive mammogram compared to having a normal one.

Competing interests None.

Correction notice This article has been corrected since it was published Online First. Reference 33 was a repeat of reference 27 and so has been deleted from the article.

References

1. **Godlee F.** Mammography wars. *BMJ* 2011;**343**:d7623.
2. **Thornton H,** Pryke M, Mackie A. Mammography wars: rapid responses. *BMJ* 2011;**343**:d7623. http://www.bmj.com/content/343/bmj.d7623?tab=responses
3. **Kopans DB,** Smith RA, Duffy SW. Mammographic screening and overdiagnosis. *Radiology* 2011;**260**:616–20.
4. **Gotzsche PC,** Nielsen M. Screening for breast cancer with mammography. *Cochrane Rev* 2011;**(7)**:CD001877.
5. **Jorgensen KJ,** Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;**339**:b2587.
6. **Seigneurin A,** Francois O, Labarere J, *et al.* Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. *BMJ* 2011;**343**:d7071.
7. **Brewer NT,** Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007;**146**:502–10.
8. **Marteau TM,** Cook R, Kidd 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.
9. **French DP,** Maissi E, Marteau TM. The psychological costs of inadequate cervical smear test results: three-month follow-up. *Psychooncology* 2006;**15**:498–508.
10. **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.
11. **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.
12. **Armstrong K,** Moye E, Williams S, *et al.* Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Ann Intern Med* 2007;**146**:516–26.
13. **Brett J,** Bankhead C, Henderson B, *et al.* The psychological impact of mammographic screening. A systematic review. *Psycho Oncol* 2005;**14**:917–38.
14. **Bankhead CR,** Brett J, Bukach C, *et al.* The impact of screening on future health-promoting behaviours and health beliefs: a systematic review. *Health Technol Assess* 2003;**7**:1–92.
15. **Armstrong K,** Moye E, Williams S, *et al.* 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.
16. **International Agency for Research on Cancer, W.H.O. (2002).** Breast Cancer Screening (Rep. No. 7), Lyon, France: IARC Press.
17. **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.
18. **Egger M,** Smith GD, Altman DG. *Systematic reviews in health care: meta-analysis in context.* 2nd edn. London: BMJ, 2001.
19. **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.
20. **Elm Ev,** Altman DG, Egger M, *et al.* Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8.
21. **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.
22. **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.
23. **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.
24. **Ellman R,** Angeli N, Christians A, *et al.* Maguire P. Psychiatric morbidity associated with screening for breast cancer. *Br J Cancer* 1989;**60**:781–4.
25. **Bull AR,** Campbell MJ. Assessment of the psychological impact of a breast screening programme. *Br J Radiol* 1991;**64**:510–15.
26. **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.
27. **O'Sullivan I,** Sutton S, Dixon S, *et al.* False positive results do not have a negative effect on reattendance for subsequent breast screening. *J Med Screen* 2001;**8**:145–8.
28. **Orton M.** Factors affecting women's response to an invitation to attend for a second breast cancer screening examination. *Br J Gen Pract* 1991;**349**:320–2.
29. **Meldrum P,** Turnbull D, Dobson HM, *et al.* Tailored written invitations for second round breast cancer screening: a randomised controlled trial. *J Med Screen* 1994;**1**:245–8.
30. **Cockburn J,** De Luise T, Hurley S, *et al.* Development and validation of the PCQ: a questionnaire to measure the psychological consequences of screening mammography. *Soc Sci Med* 1992;**34**:1129–34.
31. **Zigmond AS,** Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;**67**:361–70.
32. **Goldberg DP.** *Manual of the general health questionnaire.* Windsor, UK: NFER-Nelson Publishing, 1978.

Appendix 1: Protocol

Technology Assessment Report commissioned by the NETSCC HTA Programme

HTA 09/145

April 2011

1. Title: The psychological consequences of false positive mammograms

2. Tar team: PenTAG, University of Exeter

Project Lead: Mary Bond
Research Fellow in HTA
PenTAG
Peninsula Medical School
Veysey Building
Salmon Pool Lane
Exeter, EX2 4SG

Telephone: 01392 726077

Fax: 01392 421009

Email: mary.bond@pms.ac.uk

3. Plain English Summary

In the UK women aged 50 to 70 years old are invited to come for mammography screening every three 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.

4. 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 three 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.[1]

Quantitative observational studies looking at the psychological and behavioural consequences of false-positive mammograms show conflicting results. Some studies indicate that, whilst 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.[2-4] Other studies report that there are long-term adverse psychological consequences to receiving a false-positive mammogram.[5-8] The outcomes of studies looking at whether having false-positive results affects future attendance at breast screening appointments is similarly conflicted.[7;9-11]

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.[12] 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.[12] 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?

    1.1. 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, sub-group 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, sub-group 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.

5. 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[13] including those for non-randomized and qualitative studies.[14]

5.1. Inclusion criteria

Question	Criteria	Specification	Notes
1 and 2	Population	Women who have received a false-positive result from routine mammogram screening in the UK and have been invited for further assessment	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
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 papes is known.

5.2. 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.

5.3. 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, 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.

5.4. 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.

5.5. 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.

5.6. 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.[13;15]

Qualitative studies will have their quality assessed using a standard assessment tool, e.g. Mays and Pope 1995[16] and Popay and colleagues 1998[17], a number of these will be piloted to assess their suitability for the task.

6. Methods for analysis and synthesis of evidence of clinical effectiveness

6.1. 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.[13;15] The quality of systematic reviews was evaluated using the PRISMA statement,[18]. Individual RCTs were appraised with the CONSORT statement[19] and individual observational studies with STROBE guidelines.[20]

6.2. Qualitative analysis and synthesis

These studies will be analysed using meta-ethnography[21-23] supported by Atlas.ti6 software. Here the included studies' results are translated into one another, whilst 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.

6.3. Combined synthesis of quantitative and qualitative evidence

The results of the quantitative and qualitative analyses will undergo narrative synthesis to construct an explanatory framework.[24;25] 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.

7. Expertise in this TAR team

7.1. 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 a 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.

7.2. 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.

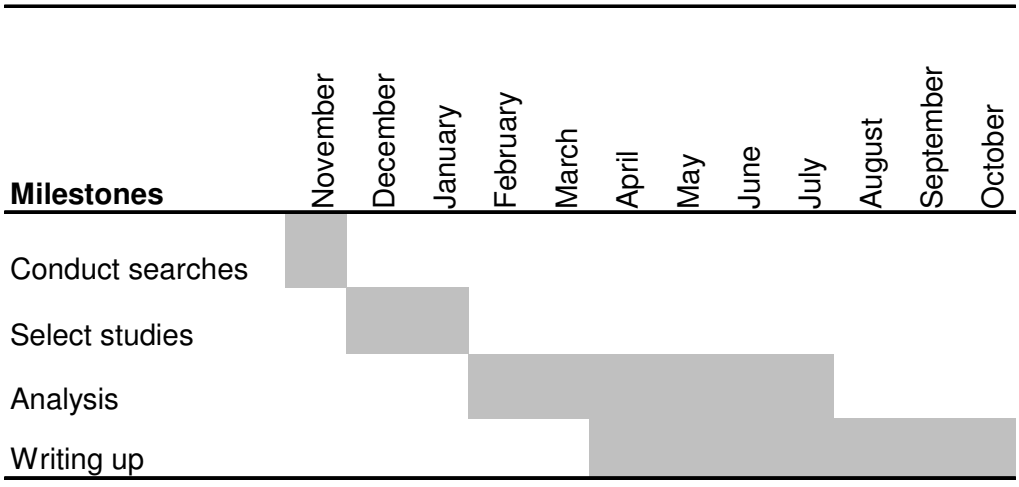
7.3. Contributions of team members

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.

8. Competing interests of authors

None.

9. Timetable and project milestones



10. Reference List

(1) NHS Breast Cancer Screening Programme. Breast Screening Results from the NHSBSP 2007/2008. http://www cancerscreening nhs uk/breastscreen/statistics html [2008 [cited 2010 July 20]; Available from: URL:http://www.cancerscreening.nhs.uk/breastscreen/statistics.html

(2) 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(4):463-469.

(3) 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(6):705-710.

(4) 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. *Journal of epidemiology and community health* 1995; 49:413-418.

(5) 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(9):1089-1097.

(6) Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. *Br J Cancer* 1990; 62(6):1018-1022.

(7) 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(4):292-300.

(8) 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(4):396-403.

(9) 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):1-6.

(10) 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(12):1730-1737.

(11) McCann J, Stockton D, Godward S. Impact of false-positive mammography on subsequent screening attendance and risk of cancer. *Breast Cancer Res* 2002; 4(5):R11.

(12) Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007; 146(7):502-510.

(13) 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.

(14) 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):iii-173.

(15) Egger M, Smith GD, Altman DG. Systematic Reviews in Health Care: Meta-analysis in Context. 2nd ed. London: BMJ; 2001.

(16) Mays N, Pope C. Qualitative Research: Rigour and qualitative research. *BMJ* 1995; 311(6997):109-112.

(17) 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(3):341-351.

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

(19) Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet* 2001; 357(9263):1191-1194.

(20) 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(7624):806-808.

(21) Noblit GW, Hare RD. Meta-ethnography: synthesizing qualitative studies. Newbury Park: Sage; 1988.

(22) 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(4):209-215.

(23) 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(6):550-562.

(24) 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. http://evi sagepub com [2009 [cited 2009 June 12]; 15(1):[49-74]

(25) 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 Education Research* 2008; 23(5):770-790.

Appendix 2: Search Strategy

This is the original Medline search strategy by KW. Other search strategies are available from the authors on request.

Databases, Host Date Searched, Years	Search Strategy Keywords added to Refman	Number of Results
Medline Ovid Scoping Search 1950- current Searched on 08/10/2010	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/	559

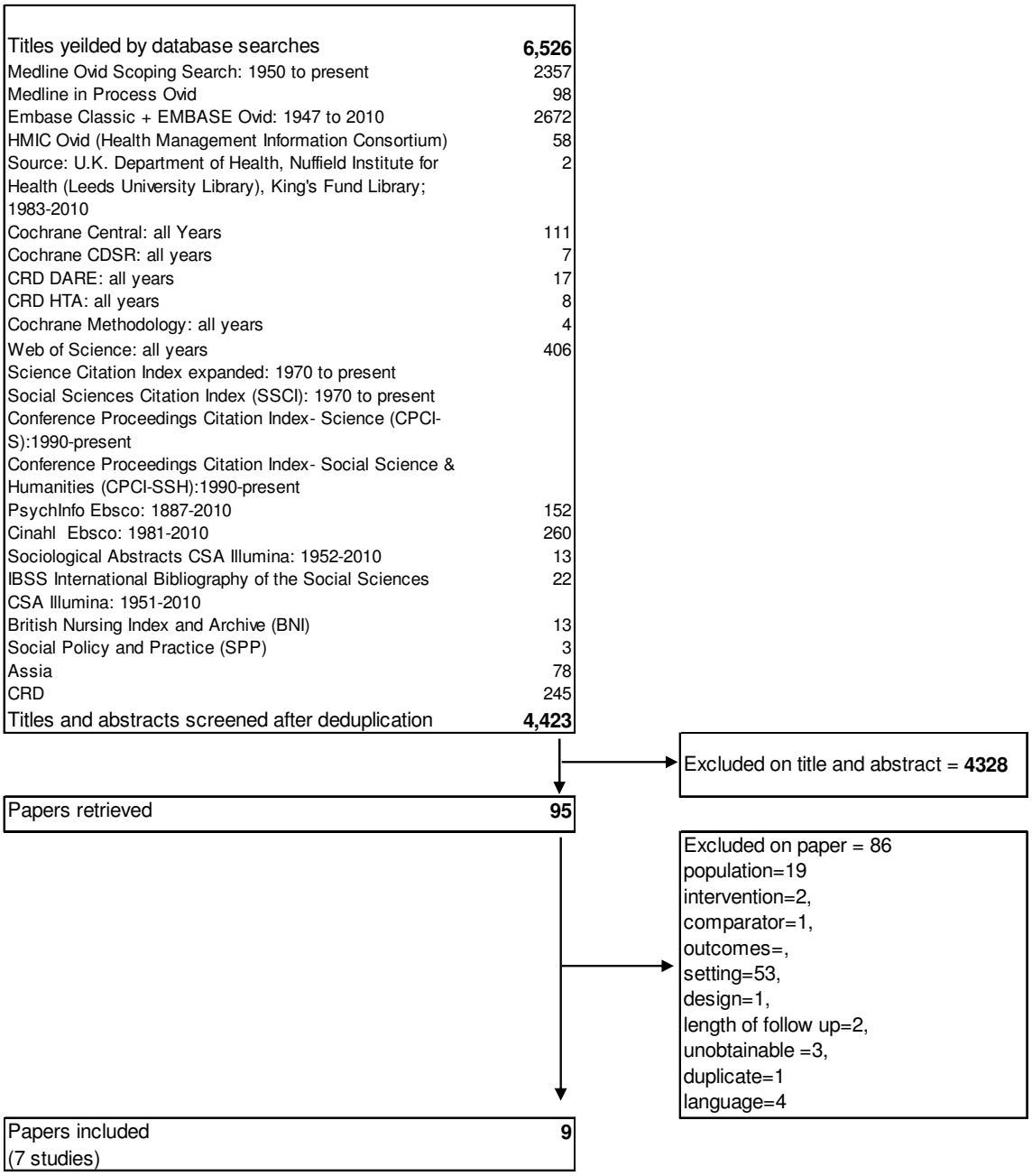
	32. exp Behavior/ 33. exp attitude/ 34. Motivation/ 35. Decision Making/ 36. exp "Quality of Life"/ 37. Health Knowledge, Attitudes, Practice/ or Attitude to Health/ or Patient Satisfaction/ or Patient Participation/ or Consumer Participation/ or Consumer Satisfaction/ or Sick Role/ or "Patient Acceptance 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. 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 counseling/ 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. 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	

Appendix 3: Study flow chart



Appendix 4:Critical appraisal of included studies

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	1	2	3	4	5	6	7	8
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 prespecified 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	✓
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	✓	✓	✓	✓	✓	✓	✓	✓
		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
		Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	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
		Case-control study—For matched studies, give matching criteria and the number of controls per case	NA	NA	NA	NA	NA	NA	NA	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	✓	✓	✓	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	✓	✓	✓	✓	✓	✓	✓	✓
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	✓	✓	✓	✓	✓	✗	✗	✓
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
		(*c*) Explain how missing data were addressed	✗	✗	✓	✗	✗	NA	NA	✗
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed	✗	✗	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
		Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	NA	NA	NA	NA	NA	NA	NA	✗
		(*e*) Describe any sensitivity analyses	NA	NA	NA	NA	NA	NA	NA	NA

1: Brett & Austoker (2001), 2:Brett et al. (1998), 3:Ong et al. (1997)a, 4:Bull & Campbell (1991), 5: Ellman et al. (1989), 6:McCann et al. (2002), 7:O'Sullivan (2001), 8:Orton et al. (1991). ✓: item present, ✗:item absent, P: item partially present, F: results only present as figures, NA: not applicable

Results			**1**	**2**	**3**	**5**	**6**	**9**	**10**	**11**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg 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	✗
		(c) Consider use of a flow diagram	✗	✗	NA	✗	✗	✗	✗	✗
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	✓	✗	P	P	P	P	✗	✗
		(b) Indicate number of participants with missing data for each variable of interest	✗	✗	✓	✗	✗	NA	NA	✗
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	✓	✓	NA	✗	✗	NA	NA	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	✓	✓	NA	✓	✓	✓	✓	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	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	✓
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	✓	✓	✓	✓	✓	✓	✓	P
		(*b*) Report category boundaries when continuous variables were categorized	NA	NA	NA	✓	✓	NA	NA	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	✗	✗	NA	NA	NA	NA	NA	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	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	✗	✓	✗	✗	✗	✗	✗	✗
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	✓	✓	✓
Generalisability	21	Discuss the generalisability (external validity) of the study results	✗	✗	✗	✗	✗	✗	✗	✓
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	✓	✓	✓	✗	✓	✓	✓	✗

1: Brett & Austoker (2001), 2:Brett et al. (1998), 3:Ong et al. (1997)a, 4:Bull & Campbell (1991), 5: Ellman et al. (1989),6:McCann et al. (2002), 7:O'Sullivan (2001), 8:Orton et al. (1991). ✓: item present, ✗:item absent, P: item partially present, F: results only present as figures, NA: not applicable

✓	**item present**
✗	**item absent**
P	**item partially present**
F	**results only presented as Figures**
NA	**not applicable**

Meldrum et al. 1994. CONSORT statement			
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 (for example, participants, care providers, those assessing outcomes) and how
	11b	Yes	If relevant, description of the similarity of interventions
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	F	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 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% confidence interval)
	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