

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 Dean¹ and C Hyde¹

¹Peninsula Technology Assessment Group (PenTAG), University of Exeter, Exeter, UK

²Karen Welch Information Consultancy, Fareham, Hampshire, UK

*Corresponding author



Executive summary

Psychological consequences of false-positive screening mammograms

Health Technology Assessment 2013; Vol. 17: No. 13

DOI: 10.3310/hta17130

NIHR Journals Library

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.

Publication

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

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 research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA programme findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

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.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Bond *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

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

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Dr Andree Le May Chair of NIHR Journals Library Editorial Group

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Health Sciences, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Tom Marshall Reader in Primary Care, School of Health and Population Sciences, University of Birmingham, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Honorary Professor, Business School, Winchester University and Medical School, University of Warwick, UK

Professor Jane Norman Professor of Maternal and Fetal Health, University of Edinburgh, UK

Professor John Powell Senior Clinical Researcher, Department of Primary Care, University of Oxford, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professorial Research Associate, University College London, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Editorial contact: nihredit@southampton.ac.uk