Pergamon

European Journal of Cancer, Vol. 34, No. 13, pp. 2010–2014, 1998

Printed in Great Britain
0959-8049/98/$—see front matter

PII: S0959-8049(98)00294-9

Original Paper

Breast Screening: the Psychological Sequelae of False-Positive Recall in Women with and without a Family History of Breast Cancer

F.J. Gilbert,[1] C.M. Cordiner,[2] I.R. Affleck,[3] D.B. Hood,[4] D. Mathieson[5] and L.G. Walker[2]

[1]Department of Radiology, University of Aberdeen Medical School; [2]Behavioural Oncology Unit, Department of Mental Health and Surgery, University of Aberdeen Medical School, Foresterhill, Aberdeen AB25 2ZD; [3]Health Centre, Aberdeen; [4]Inverurie Health Centre, Inverurie, Aberdeenshire; and [5]Cults Medical Group, Cults, Aberdeen, U.K.

The psychological effects of false-positive mammography were evaluated in 124 women who had taken part in the U.K. National Health Service Breast Screening Programme. In addition, the effects of recall on women with and without a family history were compared. These women were asked to complete the Hospital Anxiety and Depression Scale (HADS) before being invited to attend for screening, at recall and 5 weeks and 4 months after recall. At screening and at recall, the women were asked to complete the Health Questionnaire (HQ) which measures stress-related behaviour changes in the previous week. In the week before screening, compared with women who did not have a family history of breast cancer, women with a family history had lower scores on HADS depression and reported fewer stress-related behaviour changes. At recall, regardless of family history, the women were more likely to have borderline or clinically significant anxiety than at baseline or screening. Nevertheless, for most women, recall-induced anxiety was relatively transient (less than 5 weeks). Compared with women without a family history, women with a family history were more anxious 4 months after recall, although their anxiety scores tended to be lower ($P < 0.06$) than at baseline. A strength of the present study is that the initial baseline measure was uncontaminated by the screening process. Women who did not complete questionnaires at one or more of the subsequent time points scored higher on HADS depression at baseline, indicating that the results are likely to have underestimated the effects of recall. Screening appears to be less stressful for women with a family history than for those without a history. However, for both groups recall causes short term distress. Breast screening programmes should ensure that steps are taken to minimise the number of women who are recalled for unnecessary investigations. © **1998 Elsevier Science Ltd. All rights reserved.**

Key words: breast screening, psychological distress, breast cancer, familial risk
Eur J Cancer, Vol. 34, No. 13, pp. 2010–2014, 1998

INTRODUCTION

CONCERN HAS been expressed about the psychological effects of being recalled for assessment following false-positive mammographic screening [1, 2]. Citing evidence of false diagnosis from other screening programmes (including hypertension, alpha-fetoprotein levels and congenital hypothyroidism), Marteau [3] suggested that the ''psychological costs of screening may sometimes be bad enough to undermine the benefits of screening''.

Several studies have examined the psychological effects of being recalled for assessment and subsequently being found to have normal breasts or benign disease (i.e. a false-positive result) [4–9]. However, only one of these studies obtained a baseline measure of distress uncontaminated by the screening process and the authors emphasise that this study was not designed primarily to evaluate the effects of recall [9].

Correspondence to L.G. Walker.
Received 31 Mar. 1998; revised 13 Jul. 1998; accepted 23 Jul. 1998.

A neglected aspect of screening concerns the psychological effects in women who have a family history of breast cancer. It may be that they are particularly adversely affected by a false-positive recall, especially if their attention is drawn to their family history at the time of screening.

The aims of this study, therefore, were to evaluate the psychological effects of false-positive mammography and to compare the effects of recall in women with, and without, a family history of breast cancer. It was hypothesised that (a) false-positive recall will be stressful as demonstrated by elevated scores on the Hospital Anxiety and Depression Scale HADS) [10] and the Health Questionnaire (HQ) [11]; (b) stress levels will return to baseline levels during the follow-up period; (c) women with a family history will be more distressed by screening than those with no family history; and (d) compared with women without a self-reported family history, women with a family history will be more distressed by a false-positive result.

PATIENTS AND METHODS

Three health centres in North-East Scotland which were about to participate in the U.K. National Health Service Breast Screening Programme were selected. These three practices, cumulatively, had 2,357 women aged between 50 and 64 years of age.

In order to obtain a baseline that was uncontaminated by the knowledge of impending breast screening, a letter was sent out to each patient from her GP explaining that a survey of women's health was being carried out and asking her to complete the HADS [11]. A reminder was sent 2 weeks later to all non-respondents. The initial letter was sent to the women 6 weeks before they were sent an invitation to attend for breast screening.

When the women attended breast screening, they were asked to complete a second HADS immediately prior to mammography (in a few cases, for operational reasons, the HADS was completed immediately following mammography). Women who were recalled for further assessment were asked to complete another HADS at the clinic and, again, by post, 5 weeks and 4 months later. To measure stress-related behaviour changes, women were invited to complete the Health Questionnaire (HQ) at screening and at recall: because the HQ is designed to assess event related changes, it was not administered at baseline or at the two follow-ups.

HADS is a 14-item self-report questionnaire with separate measures of anxiety and depression. It is possible to analyse raw scores (the maximum worst score for each scale is 21). Alternatively, scores can be categorised into 'normal' (0–7), 'borderline' (8–10) and 'significant' (> 10) [10]. Respondents are asked to complete the test with reference to the previous 7 days.

HQ is a 7-item self-report questionnaire which was designed to measure perceived changes in various stress related behaviours and emotions during the 7 days prior to a potentially stressful event such as screening or recall (sleep, appetite, irritability, happiness, worry, concentration and ability to relax). Details and psychometric properties have been previously reported [11]. Responses to individual items (excluding appetite) can be summated to give a total score.

When the women attended for screening, the radiographer enquired about a family history of breast cancer. The radiologist subsequently evaluated the family history to determine whether the woman was likely to be at least twice the population risk of breast cancer (using criteria provided by the local Department of Medical Genetics). Women who met these criteria were recalled to the assessment clinic and informed that they were probably at increased risk. They were offered referral to the Department of Medical Genetics for detailed risk assessment and counselling. If detailed risk assessment subsequently confirmed that their risk was at least double the population risk, they were offered early recall (screening every 18 months as opposed to every 3 years).

The overall effect of assessment occasion on HADS scores and HQ total score were analysed using repeated measures multivariate analysis of variance. The effect of a family history was assessed initially by means of repeated measures multivariate analysis of variance. Between group comparisons at specified time-points were made by one-way analysis of variance (HQ scores) or covariance (HADS scores) [12]. Paired within-group comparisons were evaluated using student's t test. HADS depression scores were transformed to approximate normality using a $\log_{10}$ transformation ($\log_{10}$, score +1). Categorised HADS data were analysed using the McNemar test for the Significance of Change and between-group differences were assessed using the Pearson Chi-squared test.

Statistical tests were carried out using all available data: graphs were drawn for respondents who had complete data for all five time-points. All analyses were performed using SPSS MS for Windows. Alpha was set at 0.05 (two-tailed).

RESULTS

2,110 of the 2,357 eligible women responded to the baseline HADS sent by the general practitioner prior to the invitation to attend screening (a response rate of 90%). 1,463 (70%) women satisfactorily completed the HADS when they attended for screening and 1,561 (66%) completed the first HQ at this time.

163 of the 2,357 (7%) women reported a family history of breast cancer, although in 127 cases this was not considered significant and the women were not recalled. Of the remaining 36 who were recalled the reasons are shown in Table 1.

Altogether, 133 of the cohort were recalled to the assessment clinic. 9 of these 133 women were found to have cancer and, therefore, were excluded from the study.

Baseline HADS data were available for 122 of the 124 women (35 with a family history and 87 without family history). 90 women completed all items of the HADS on all 5 occasions and 105 women completed all items of the HQ at screening and recall.

Overall results

Repeated measures analysis of variance revealed a significant effect for assessment occasions for HADS anxiety ($F = 3.74$, $P = 0.007$) and HQ scores ($F = 16.73$, $P < 0.0005$) but not for HADS depression ($F = 2.36$, $P < 0.06$). The mean

Table 1. *Family history and recall*

	Patient *n*
Positive response to family history question at screening	163
Family history non significant, therefore, not recalled	127
Recalled because of family history and mammographic abnormality	18
Recalled because of a significant family history	17
Declined to attend recall for family history	1

anxiety and depression scores (with SEM bars) for each occasion are shown in Figures 1(a) and (b): the mean total scores (with SEM bars) from the HQ are shown in Figure 1(c).

Compared with baseline, women scored lower on HADS anxiety at final follow-up ($t = 2.70$, $P = 0.008$). The increase from baseline at recall was not statistically significant ($t = 1.73$, $P < 0.09$) and at 4-week follow-up scores were similar to those at baseline ($t = 0.32$, $P = 0.75$). Compared with screening, the women scored higher at recall ($t = 2.75$, $P = 0.007$) and lower at final follow-up ($t = 2.59$, $P = 0.01$). Anxiety scores fell between recall and first follow-up ($t = 3.08$, $P = 0.003$) and between recall and second follow-up ($t = 4.13$, $P < 0.0005$).

Although the repeated measures analysis of variance was not significant at the 5% level ($P < 0.06$), it is of interest to note that HADS depression scores fell significantly between baseline and screening ($t = 2.04$, $P = 0.04$). Compared with screening, scores increased at recall ($t = 2.25$, $P = 0.03$).

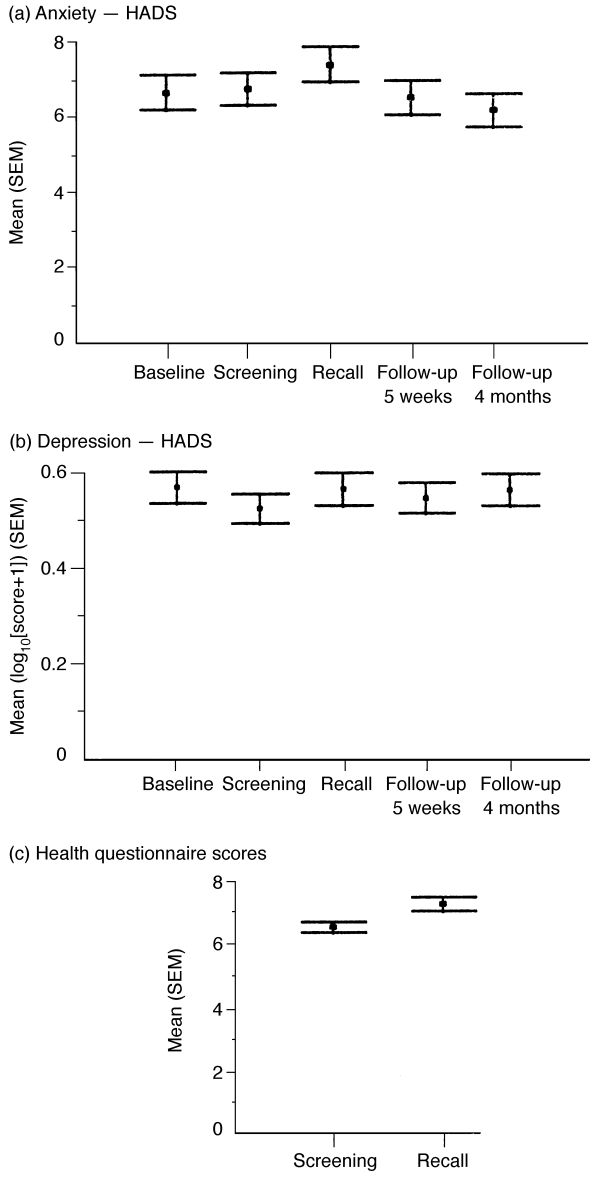

Figure 1. Hospital Anxiety and Depression Scale (HADS) and Health Questionniare (HQ) scores.

The McNemar Test for the Significance of Change was applied to HADS clinical categories (normal, 7 or less, versus borderline or significant, 8 or above). For anxiety, women were more likely to have borderline or clinically significant anxiety at recall than at baseline (binomial $P < 0.05$), screening ($P < 0.0001$), first follow-up ($P < 0.005$) or at final follow-up ($P < 0.02$). None of the comparisons was significant for depression. The HADS scores by clinical category for each assessment occasion are shown in Figures 2(a and b).

HQ scores increased significantly between screening and recall ($F = 15.6$, $P < 0.0005$).

The effect of a positive family history

Repeated measures analysis of variance for HADS anxiety showed an overall significant effect for assessment occasions ($F = 4.09$, $P < 0.004$) and family history ($F = 4.05$, $P < 0.05$). The interaction was not significant ($F = 1.07$, ns). Compared with women without a family history of breast cancer, women with a family history were significantly more anxious at the second follow-up assessment ($F = 4.14$, $P = 0.045$). However, at the second follow-up, the anxiety scores of women without a family history were significantly lower than at baseline ($F = 4.57$, $P = 0.05$) and there was a tendency for this also to be the case for women with a family history ($F = 3.61$,

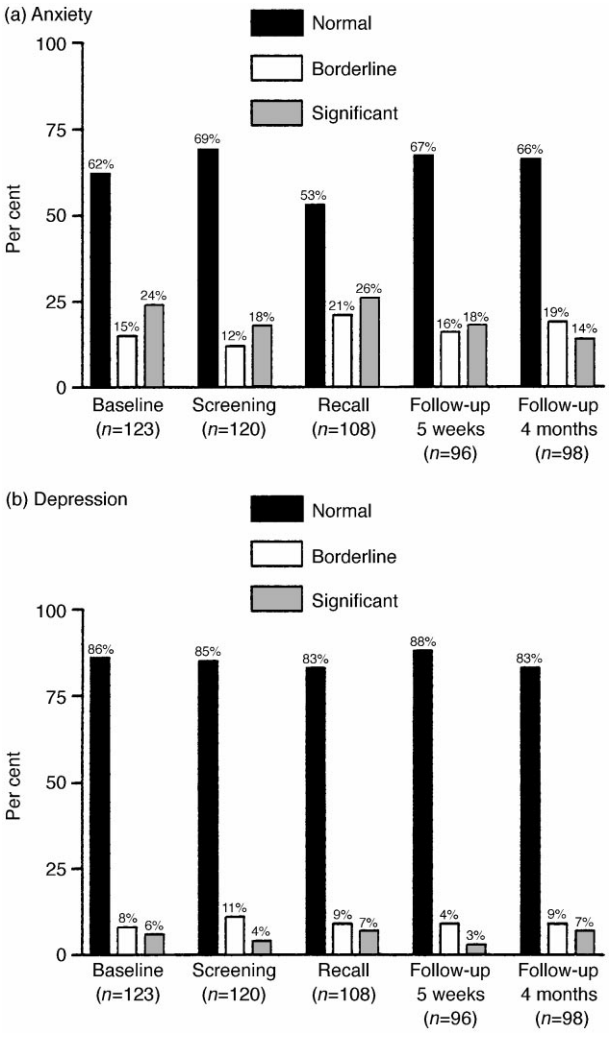

Figure 2. HADS scores by clinical categories.

$P = 0.06$). When the data were analysed by clinical categories, there were no significant differences in the proportions of women with borderline or significant anxiety at any time point.

Repeated measures analysis of variance for depression scores was significant for the interaction between family history and assessment occasions ($F = 2.51$, $P = 0.05$). However, none of the one-way ANCOVAs (baseline score as covariate) was significant. Women with a family history were less likely to score in the borderline or significant range of depression at screening than those with no family history ($\chi^2 = 5.76$, $P < 0.02$).

A repeated measures between-group analysis of variance of HQ scores was not significant ($F = 3.15$, $P < 0.08$). However, women with a family history reported less stress-related changes than those with a family history when they attended for screening ($F = 6.38$, $P = 0.01$) but not at recall ($F = 0.05$, ns).

Characteristics of patients with incomplete data

Women who did not return the first and second follow-up questionnaires had significantly higher baseline HADS depression scores than those who did return these questionnaires ($t = 2.61$, $P = 0.01$ and $t = 1.96$, $P = 0.05$, respectively). Compared with scores at screening, women who did not complete questionnaires at recall tended to score higher on HADS depression than those who did complete questionnaires at recall ($t = 1.89$, $P < 0.06$). Similarly, women who did not return questionnaires at the first or second follow-up scored higher on HADS depression at screening than those who completed questionnaires at these time points ($t = 2.46$, $P = 0.02$ and $t = 2.26$, $P = 0.03$, respectively). Finally, women who did not complete the second follow-up questionnaires were more depressed ($t = 1.96$, $P = 0.05$) and anxious ($t = 1.99$, $P < 0.06$) at recall than those who had completed the second follow-up questionnaires.

Effects of repeat administration of the questionnaires. To evaluate the possibility that repeated administration of the questionnaires may have influenced the findings, a small subsidiary study was subsequently carried out. A random sample of 100 women from one of the practices was identified. They were all between 50 and 64 years of age. As in the main study, they were sent a letter from their general practitioner asking them to complete a HADS and an HQ. 6 weeks later, they were sent a second HADS and HQ. 84 women satisfactorily completed the initial questionnaires and 82 returned the second set. Mean scores (first assessment versus second assessment) were as follows: HADS anxiety 5.66 versus 5.61 ($t = 0.18$, $P = 0.85$), HADS depression 3.66 versus 3.54 ($t = 0.46$, $P = 0.65$), HQ total score 6.71 versus 6.72 ($t = 0.04$, $P = 0.97$). However, it is important, to note that the correlations between first and second assessments were moderate for HADS anxiety ($r = 0.63$, $P < 0.0005$) and HADS depression ($r = 0.69$, $P < 0.0005$) and non-significant for HQ total score ($r = -0.08$, $P = 0.43$). These data indicate that, in this sample, mean scores are remarkably stable over a period of 6 weeks; however, individual scores change considerably during this time as should be the case for state measures.

DISCUSSION

The present study has a number of methodological strengths, including the use of a baseline psychological assessment made before the women knew that they were to be invited for breast screening, serial measurements of the women over a period of approximately 6 months, a very large initial cohort (2,357 women) and the use of questionnaires which yield information about the likely clinical significance of changes in mood (HADS) and self-perceived changes in stress related behaviours (HQ). In addition, a subsidiary study indicated that it is very unlikely that repeated administration of the questionnaire had a significant effect on the results of the main study.

The results showed that women were more likely to have borderline or clinically significant anxiety at recall than at baseline or at screening. Compared with recall, mean anxiety scores were lower at screening but not at baseline. Women were not more depressed at recall than at baseline as assessed by mean scores and clinical categories, although mean scores were higher at recall than at screening. HQ scores indicated significant stress related behaviour changes in the week prior to recall compared with the week prior to screening. These findings are consistent with previous studies [5, 6].

However, these data also indicate that for this group of women as a whole the stress caused by recall was transient. Compared with mean anxiety scores at screening, mean anxiety scores had fallen significantly by 5 weeks after recall and remained significantly lower at 4-month follow-up. Scores were also significantly lower at 4-month follow-up than at baseline. Similarly, the proportion of women having borderline or clinically significant anxiety at either follow-up did not differ significantly from baseline. Depression scores and HQ scores were similar to baseline at the two follow-up points. Cockburn and colleagues [5] assessed emotional and physical dysfunction 1 week after notification of a normal result following recall and 8 months thereafter. They found that dysfunction remained high 1 week following the normal result, but, it had returned to normal by 8 months. Ellman and colleagues [6] found no evidence for sustained anxiety 3 months after recall as assessed by the General Health Questionnaire. The present findings indicate that, overall, distress levels return to normal within 5 weeks. Given the sample size, however, the possibility that there is a small sub-group of women who have persistent severe distress cannot be excluded.

Contrary to the hypothesis, women with a family history were more likely to score in the normal range of depression at screening. Also, they reported fewer stress-induced behaviour changes in the week prior to screening. However, the data for anxiety revealed no significant differences between the two groups. Screening appears to be reassuring for women with a family history. As hypothesised, at the second follow-up point, women with a family history were significantly more anxious than those without a family history. However, their scores at the second follow-up were lower than their baseline scores and this almost reached statistical significance ($P < 0.06$). There is no evidence, therefore, of sustained anxiety in women with a family history.

Bull and Campbell [4] criticised the use of the HADS in the context of screening as they considered the tool lacked adequate sensitivity to detect screen-induced changes in mood. However, in this study the HADS was successful in demonstrating statistically and clinically significant changes even in relatively small sub-groups.

As the present study incorporated a baseline uncontaminated by the screening process, it is possible to comment on the mental health of patients who did not respond to questionnaires at subsequent time-points. Essentially, the data

indicate that women who did not complete questionnaires at one or more of the time-points following screening scored higher on HADS depression at baseline and at screening. Consequently, despite the overall good response rate, the effects of recall on depression scores are likely to have been underestimated in this study. The results need to be interpreted with this response bias in mind.

CONCLUSIONS

This study has changed our practice. Women with a family history are no longer recalled to assessment clinics as we have shown this causes raised anxiety despite informing them that their mammogram is normal. These women are sent their results as part of the normal screening programme and are then given the opportunity to attend a genetics clinic. Approximately 50% take up this offer.

In conclusion, recall because of a false-positive mammogram causes significant adverse psychological effects. However, for women without a family history of breast cancer, these appear to be transient (less than 5 weeks). Breast screening programmes should ensure that steps are taken to minimise the number of women who are recalled for unnecessary investigations.

1. Forrest APM. *Breast Cancer Screening: Report to the Health Ministers of England, Wales, Scotland and Northern Ireland.* London, HMSO, 1986.
2. Maguire GP. Possible complications of screening for breast cancer. *Br J Radiol* 1982, **56**, 284.
3. Marteau TM. Reducing the psychological costs of screening. *Br Med J* 1989, **299**, 527–528.
4. Bull AR, Campbell MJ. Assessment of the psychological impact of a breast screening programme. *Br J Radiol* 1991, **64**, 510–515.
5. Cockburn J, Staples M, Hurley SF, De Luise T. Psychological consequences of screening mammography. *J Med Screen* 1994, **1**, 7–12.
6. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. *Br J Cancer* 1989, **60**, 781–784.
7. Gram IT, Slenker SE. Cancer anxiety and attitudes toward mammography among screening attenders, non-attenders and women never invited. *Am J Pub Health* 1992, **82**, 249–251.
8. Lerman C, Trock B, Rimer BK, *et al.* Psychological and behavioural implications of abnormal mammograms. *Ann Intern Med* 1991, **114**, 657–661.
9. Sutton S, Saidi G, Bickler G, Hunter J. Does routine screening for breast cancer raise anxiety? Results from a three wave prospective study in England. *J Epidem Com Health* 1995, **49**, 413–418.
10. Zigmond AS, Snaith RT. The Hospital Anxiety and Depression Scale. *Acta Psychiat Scand* 1983, **67**, 361–337.
11. Walker LG, Cordiner C, Gilbert FJ, *et al.* How distressing is attendance for routine breast screening. *Psycho-Oncol* 1994, **3**, 299–304.
12. Tabachnik BG, Fidell LS. *Using Multivariate Statistics* (3rd Edition). New York, Harper-Collins, 1996.

Acknowledgements—We gratefully acknowledge the help of many people including the women who completed the questionnaires, the general practitioners and other staff from the Health Centres at Cults, 526 King Street, Aberdeen, and Inverurie, Mrs E. Singleton and colleagues at the Breast Screening Centre, Ms J. Anderson and Miss S. McLeod. Funding was generously provided by the Acceptability Committee of the Scottish Breast Screening Programme and Grampian Health Board (Endowments).