

## Comparing hospital and telephone follow-up after treatment for breast cancer: randomised equivalence trial

Kinta Beaver, professor of nursing,<sup>1</sup> Debbie Tysver-Robinson, nurse consultant,<sup>2</sup> Malcolm Campbell, lecturer in statistics,<sup>1</sup> Mary Twomey, research associate,<sup>1</sup> Susan Williamson, research fellow,<sup>1</sup> Andrew Hindley, consultant clinical oncologist,<sup>3</sup> Shabbir Susnerwala, consultant clinical oncologist,<sup>3</sup> Graham Dunn, professor of biomedical statistics,<sup>4</sup> Karen Luker, professor of nursing<sup>1</sup>

<sup>1</sup>School of Nursing, Midwifery and Social Work, University of Manchester, Manchester M13 9PL

<sup>2</sup>Blackpool, Fylde and Wyre Hospitals NHS Foundation Trust, Blackpool, Lancashire

<sup>3</sup>Rosemere Cancer Centre, Royal Preston Hospital, Preston, Lancashire

<sup>4</sup>Health Methodology Research Group, University of Manchester, Manchester

Correspondence to: K Beaver  
kinta.beaver@manchester.ac.uk

Cite this as: *BMJ* 2009;338:a3147  
doi:10.1136/bmj.a3147

### ABSTRACT

**Objective** To compare traditional hospital follow-up with telephone follow-up by specialist nurses after treatment for breast cancer.

**Design** A two centre randomised equivalence trial in which women remained in the study for a mean of 24 months.

**Setting** Outpatient clinics in two NHS hospital trusts in the north west of England

**Participants** 374 women treated for breast cancer who were at low to moderate risk of recurrence.

**Interventions** Participants were randomised to traditional hospital follow-up (consultation, clinical examination, and mammography as per hospital policy) or telephone follow-up by specialist nurses (consultation with structured intervention and mammography according to hospital policy).

**Main outcome measures** Psychological morbidity (state-trait anxiety inventory, general health questionnaire (GHQ-12)), participants' needs for information, participants' satisfaction, clinical investigations ordered, and time to detection of recurrent disease.

**Results** The 95% confidence interval for difference in mean state-trait scores adjusted for treatment received (-3.33 to 2.07) was within the predefined equivalence region (-3.5 to 3.5). The women in the telephone group were no more anxious as a result of foregoing clinic examinations and face-to-face consultations and reported higher levels of satisfaction than those attending hospital clinics (intention to treat  $P < 0.001$ ). The numbers of clinical investigations ordered did not differ between groups. Recurrences were few (4.5%), with no differences between groups for time to detection (median 60.5 (range 37-131) days in hospital group v 39.0 (10-152) days in telephone group;  $P = 0.228$ ).

**Conclusions** Telephone follow-up was well received by participants, with no physical or psychological disadvantage. It is suitable for women at low to moderate risk of recurrence and those with long travelling distances or mobility problems and decreases the burden on busy hospital clinics.

**Trial registration** National Cancer Research Institute 1477.

### INTRODUCTION

In many countries clinical examination, consultation, and routine mammography form the basis of routine follow-up for women in remission from breast cancer, with the primary objective of detecting recurrent disease.<sup>1</sup> This form of surveillance is as effective in terms of overall survival and quality of life as more intensive approaches,<sup>1</sup> although whether routine follow-up translates into improved survival is questionable.<sup>2</sup> Routine mammography alone might confirm local recurrence in most treated women.<sup>3</sup> Recurrences often present as interval events and are not usually detected by clinical examination of patients without symptoms.<sup>4-8</sup> National guidelines in the United Kingdom state that intensive follow-up to detect metastatic disease is not beneficial, although patients should have continued access to specialist breast care nurses for advice and support.<sup>9</sup>

Follow-up consultations could provide an opportunity to meet information and psychosocial needs, although this might be challenging given that the mean duration for consultations is six minutes.<sup>10</sup> Alternative strategies, focusing less on survival and more on patients' satisfaction, have shown some benefit to patients. General practitioners can be as effective at providing follow-up care as hospital doctors,<sup>7</sup> but there is a clear need for appropriate training and resources to transfer follow-up care from hospital to community.<sup>11</sup> In the UK, follow-up continues primarily in hospital clinics.

We based our study on previous work in the UK and Canada,<sup>12,13</sup> which showed that women with breast cancer need specific types of information and that new follow-up strategies should be developed and evaluated.<sup>10</sup> With a proliferation of specialist nursing posts worldwide, nurse led clinics are held for different cancer groups. Breast care nurses are uniquely placed to address the information and psychosocial needs of affected women<sup>14,15</sup> and to provide follow-up services. We built on existing evidence of the effectiveness of telephone interventions for people with cancer<sup>16</sup> to deliver a structured intervention aimed primarily at meeting needs for information. Communication by

telephone is internationally relevant, particularly for people in remote areas where travel to hospital is inconvenient, time consuming, and costly.

## METHODS

In this equivalence trial<sup>17</sup> we examined whether, despite foregoing face-to-face consultations and clinical examinations, patients in the telephone arm of the study had anxiety levels that were no different from those of patients in the hospital arm.

### Study sites and sample

All participants had been treated for breast cancer at a large district general hospital or a specialist breast unit in the north west of England. Inclusion criteria included completion of primary treatment (surgery, radiotherapy, chemotherapy), no evidence of recurrent disease, low to moderate risk of recurrence,<sup>18</sup> access to a telephone, and adequate hearing. We primarily determined the risk of recurrence using the Nottingham prognostic index, considering tumour size, spread to lymph nodes, and grade of cancer. We also considered HER2 status as routine testing was being introduced at the onset of the study. Consultant surgeons and oncologists at the two study locations agreed a written protocol for determination of risk category. We included women with grade I and grade II tumours if the tumour size was  $\leq 50$  mm with three or fewer nodes affected. Women with grade III tumours were included only if they were postmenopausal,

tumour size was  $\leq 50$  mm, no nodes were affected, oestrogen receptor status was positive, and HER2 status was negative. We excluded inflammatory carcinomas and sarcomas. With these inclusion criteria no participant had a prognostic index  $>4.1$ , indicating a low to moderate risk of recurrence.

We identified consecutive eligible patients in hospital clinics, discussed the study after appointments, and subsequently contacted individuals for verbal and written consent. Women who consented were randomised to telephone or hospital follow-up. Researchers contacted a central telephone randomisation service to discover individual group allocation. Allocation sequences were computer generated with randomised permuted blocks with randomly varying block sizes, stratified by study site and whether the participant was on three, six, or 12 monthly follow-up. Allocation sequences were concealed until interventions were assigned. The analyst was blind to study group allocation. Breast care nurses had no involvement in randomisation or data collection procedures.

### Procedures and intervention

Participants randomised to the hospital group continued with hospital follow-up as per hospital policy. At the district general hospital participants were reviewed every three months for two years, six monthly for two years, then annually for a further year. At the specialist breast unit they were reviewed annually for 10 years. Preliminary work for the study had indicated that hospital consultations were generally unstructured but primarily consisted of a clinical examination, a check on whether hormone treatment was being taken as prescribed, and ordering mammograms if necessary.<sup>10</sup> As per hospital policy, both study locations allocated 10 minutes for each individual hospital appointment. Hospital consultations could be conducted by various health professionals including consultant surgeons, consultant oncologists, registrars, more junior doctors, or specialist nurses. It was more usual at both locations, however, for junior medical staff to conduct hospital appointments.

Participants randomised to telephone follow-up received telephone appointments from breast care nurses at intervals consistent with hospital follow-up policy. Appointment cards provided participants with a date and time for their appointment; all telephone appointments were registered on computerised hospital information systems so medical records staff could access patients' notes before telephone clinics. Each individual telephone appointment was allocated 30 minutes; 20 minutes for conducting the consultation and 10 minutes to dictate the outcome of consultations. Throughout the study the same specialist nurse contacted each participant in the telephone group for all appointments.

We developed a structured telephone intervention from previous findings on information needs of women with breast cancer,<sup>12,13</sup> adapting a research instrument for use as a clinical guide.<sup>19</sup> Specific questions related to changes in condition, new symptoms, and information

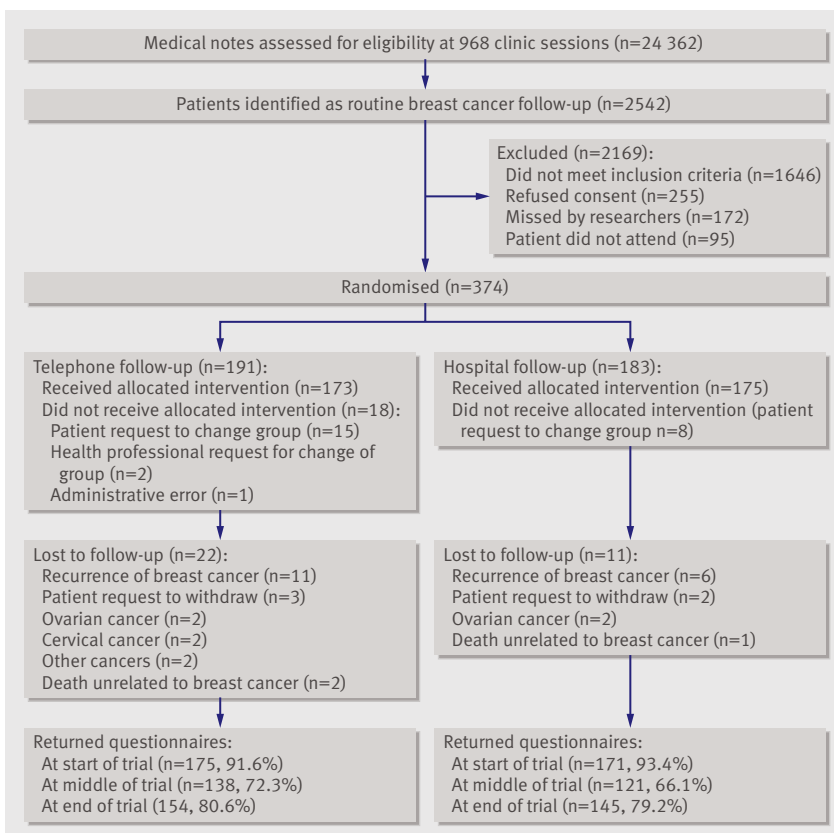


Fig 1 | Flow of participants through trial

**Table 1** Sociodemographic characteristics of women with breast cancer according to hospital or telephone follow-up. Figures are numbers (percentages) of women unless stated otherwise

	Hospital (n=183)	Telephone (n=191)	Total (n=374)
Age at recruitment (years):			
Mean (SD)	64.0 (11.1)	63.9 (10.1)	64.0 (10.6)
Median (range)	63.0 (36-87)	64.0 (42-93)	64.0 (36-93)
<45	9 (5)	4 (2)	13 (4)
45-54	26 (14)	32 (17)	58 (16)
55-64	67 (37)	60 (31)	127 (34)
65-74	46 (25)	60 (31)	106 (28)
≥75	35 (19)	35 (18)	70 (19)
Marital status:			
Married/cohabiting	117 (64)	123 (64)	240 (64)
Divorced/separated	15 (8)	21 (11)	36 (10)
Widowed	38 (21)	39 (20)	77 (21)
Never married	13 (7)	8 (4)	21 (6)
Employment status at recruitment*:			
Employed full time	28 (16)	29 (15)	57 (15)
Employed part time	29 (16)	29 (15)	58 (16)
Retired	115 (64)	121 (64)	236 (64)
Unemployed	5 (3)	2 (1)	7 (2)
Long term sick	3 (2)	7 (4)	10 (3)
Other	1 (1)	2 (1)	3 (1)
Social class (based on current/previous occupation)*:			
Managers and senior officials	13 (8)	19 (12)	32 (10)
Professional	15 (10)	18 (11)	33 (10)
Associate professionals/technical	20 (13)	20 (12)	40 (13)
Administrative/ secretarial	48 (31)	35 (22)	83 (26)
Skilled trades	11 (7)	7 (4)	18 (6)
Personal service	9 (6)	23 (14)	32 (10)
Sales/customer services	12 (8)	13 (8)	25 (8)
Process, plant, and machine operatives	4 (3)	2 (1)	6 (2)
Elementary occupations	25 (16)	25 (15)	50 (16)

\*Small numbers of missing values.

requirements about spread of disease, treatment and side effects, genetic risk, sexual attractiveness, self care (diet, support groups, finances), and family concerns. Standard protocols related to routine mammography were unaltered.

Breast care nurses underwent four half day training sessions on the administration of the telephone intervention with subsequent feedback and debriefing sessions throughout the study period. Seven nurses received training, although one nurse at the district general hospital and three nurses at the specialist breast unit conducted most telephone appointments. To monitor the integrity of the intervention, all telephone consultations were recorded with consent of the women.

#### Outcome measures

Outcomes included psychological morbidity, information needs, participants' satisfaction, clinical investigations ordered, and time to detection of recurrent disease. We designed and piloted questionnaires on information needs and participants' satisfaction.<sup>19</sup> We measured psychological morbidity using two

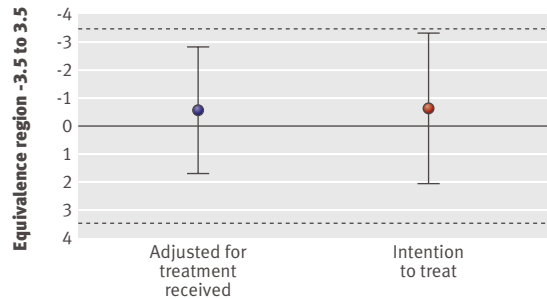
instruments. The Spielberger state-trait anxiety inventory (STAI) is a 40 item self report instrument distinguishing between short term anxiety (state) initiated by current life events (STAI Y1, 20 items) and anxiety as a personality trait (STAI Y2, 20 items, administered only initially).<sup>20</sup> The general health questionnaire (GHQ-12) is a 12 item, well validated instrument focusing on broad aspects of psychological morbidity and designed to be self administered.<sup>21</sup> We also examined cost effectiveness and will report on this elsewhere.

We collected clinical data prospectively. A record of visit form recorded actions resulting from consultations and indicators of recurrence in the hospital arm. The recorded telephone appointments provided equivalent data. We used two key index dates to establish time to diagnosis of recurrence: date of first presentation of symptoms (or indicator of recurrence) and date recurrence was confirmed to participants. Time to diagnosis of recurrence was taken as the difference between the two index dates. As the sample included only those at low to moderate risk of recurrence we expected only a few recurrences during the study period. We could not draw firm conclusions relating to detection of recurrence but compared recurrences between groups. At the end of the study we retrospectively examined all participants' case notes to check accuracy of data on recurrence.

For practical reasons we could not administer outcome questionnaires before randomisation. We sent initial questionnaires to patients immediately after randomisation, a minimum of three months before their next appointment. Questionnaires were also administered at the middle and end of the trial, established on an individual basis, depending on how long participants remained in the study. These were posted to participants immediately after telephone or hospital appointments to maximise participants' recall. Members of the research team posted all outcome measures with pre-paid return postal service. Health professionals who conducted follow-up consultations as part of this study had no involvement in the administration of outcome measures.

#### Statistical analysis

Contrary to a standard comparative trial, an equivalence trial has an alternative hypothesis of no difference between treatments or services.<sup>22</sup> We calculated our sample size on the basis of findings from pilot work related to state-trait scores.<sup>19</sup> The pilot study compared scores for the intervention (telephone follow-up by specialist nurses) and control group (hospital follow-up); we used differences in mean scores between groups to define an equivalence region for the current study. Our study was powered on the basis of a 95% equivalence region of -3.5 to 3.5 for the difference in mean state-trait scores between groups at the end of the trial, 3.5 being 10% of the expected control mean, with an expected SD of 10.0. The required sample size was 129 per group.<sup>23</sup> Assuming 20% drop out by the end of the trial, we aimed to achieve 162 per group.



**Fig 2** | Mean differences with 95% confidence intervals for telephone group minus hospital group for state-trait anxiety inventory according to method of analysis

We entered and analysed data with SPSS, release 15. Scoring of standard psychological measures was conducted as recommended in scoring manuals.<sup>20 21</sup> State-trait scores range from 20 to 80 and GHQ-12 scores from 0 to 12 (scores  $\geq 4$  indicate caseness), with higher scores on both measures indicating higher levels of psychological morbidity. We compared the outcome used to assess equivalence (the state-trait score) between groups in two ways, starting with an intention to treat analysis. When there are protocol violations, the intention to treat effect will be biased in favour of equivalence.<sup>24</sup> We used a two stage approach to estimate the effects of treatment received adjusted for

**Table 2** | Details of disease and treatment in women with breast cancer according to hospital or telephone follow-up. Figures are numbers (percentages) of women

	Hospital (n=183)	Telephone (n=191)	Total (n=374)
How breast cancer was detected:			
Breast screening programme	78 (43)	83 (44)	161 (43)
Patient presented with symptoms	99 (54)	107 (56)	206 (55)
Other	6 (3)	1 (0.5)	7 (2)
Type of surgery:			
Mastectomy	65 (36)	78 (41)	143 (38)
Lumpectomy/wide local excision	118 (65)	113 (59)	231 (62)
Received radiotherapy*	109 (60)	106 (56)	215 (58)
Received chemotherapy*	18 (10)	16 (8)	34 (9)
Time from diagnosis† (months):			
Median (range)	20.0 (1-107)	18.0 (1.25-92)	20.0 (1-107)
≤12	56 (31)	70 (37)	126 (34)
13-24	57 (31)	51 (27)	108 (29)
25-36	26 (14)	24 (13)	50 (13)
>36	43 (24)	46 (24)	89 (24)
Time from first visit after treatment‡ (months):			
Median (range)	17.0 (0.5-106)	15.0 (0.5-91)	15.0 (0.5-106)
≤12	75 (41)	87 (46)	162 (43)
13-24	45 (25)	39 (20)	84 (23)
25-36	36 (20)	27 (14)	63 (17)
>36	26 (14)	38 (20)	64 (17)
Follow-up status:			
3 monthly	61 (33)	67 (35)	128 (34)
6 monthly	71 (39)	73 (38)	144 (39)
12 monthly	51 (28)	51 (27)	102 (27)

\*Small numbers of missing values.

†One value missing because participant was initially diagnosed and treated in different area; no accurate data available.

the hidden effects of non-compliance—an adjusted treatment received analysis.<sup>25</sup> We used regression, firstly, to predict treatment received from randomised group and pre-randomisation covariates and, secondly, to predict the outcome variable from treatment received, the residuals from the first stage and the same covariates.<sup>26</sup> This is equivalent to a two stage least squares approach.<sup>27</sup> Regression models for the first stage of analyses adjusted for treatment received were fitted as general linear models. Models for intention to treat analyses and the second stage of analyses adjusted for treatment received across the three time points were fitted with linear mixed models, with either randomisation group or treatment received, respectively, stratification variables (site and follow-up status), and time point as fixed factors and patient as a random factor.<sup>28</sup> Such models allow for non-independent observations and use all outcome data (we dropped the small number of patients with no state-trait data from the analysis), assuming missing data to be missing at random.<sup>29</sup>

We initially fitted saturated models but as no interactions were significant we have presented results for main effects models. Standard deviations of residuals were similar over combinations of fixed factors except where subgroup sizes were small. For assessing equivalence, we compared state-trait scores between study groups using a two sided 95% confidence interval for the difference between estimated marginal means. Results are also given for similar intention to treat and adjusted treatment received analyses at specific time points, fitted as general linear models. We compared other measures between randomised groups under intention to treat analysis using  $\chi^2$  test and Fisher's exact test, including its extension to tables larger than  $2 \times 2$  (nominal measures), the  $\chi^2$  test for trend (ordinal), the Mann-Whitney U test (skewed), and the unpaired *t* test (interval). Outcomes were compared between time points using the McNemar test (dichotomous), the Wilcoxon matched pairs test (skewed), and the paired *t* test (interval). Logistic regression was used to assess association between sociodemographic or clinical characteristics and participation at the end of the trial.

### Monitoring

A trial management group, including clinicians, researchers, and lay representatives, met on several occasions to discuss and debate key aspects of the study and to monitor study progress.

### RESULTS

Recruitment took place between March 2003 and August 2005. Data collection continued until October 2006; patients remained in the study for a mean of 24 months (range 2-43 months). Many patients approached who met the inclusion criteria were willing to participate (374/629, 60%, fig 1); 215 patients who did not take part provided sociodemographic and treatment information and a reason for refusal, including a preference for face-to-face consultations



and a preference for clinical examination. Some family members did not want their relative to take part as they thought they would be excluded from consultations if the appointment was conducted by telephone. Participants were randomised to hospital (183, 49%) or telephone (191, 51%) follow-up. Initially, sociodemographic and clinical characteristics for those randomised were broadly similar in the two study groups (tables 1 and 2). The sample was representative in terms of age, with more participants in the 55-64 year age group than in the other age group categories (table 1), in line with data from the UK Office for National Statistics.<sup>30</sup> Participants were a median of 20 months from diagnosis, although most (63%) were 24 months or less from diagnosis (table 2). Over the course of the study over 500 telephone and 500 hospital appointments were conducted with participants. Those who refused to take part differed from participants in study site, social class, and follow-up status. Patients at the specialist breast unit (71%) were more likely to want to participate than those at the district general hospital (61%,  $\chi^2=5.01$ ,  $df=1$ ,  $P=0.025$ ), participants from higher social classes (professional occupations) were more likely to want to participate than those from lower social classes ( $\chi^2=15.77$ ,  $df=8$ ,  $P=0.046$ ), and

participants with three to 12 months between visits (67.7%, 70.6%) were more likely to participate than those on six monthly follow-up (58.1%,  $\chi^2=7.66$ ,  $df=2$ ,  $P=0.022$ ). Time from diagnosis did not differ significantly for those who did or did not take part ( $t=-0.26$ ,  $P=0.80$ ); those who refused to take part were a median of 21 months from diagnosis.

There were 26 protocol violations between baseline and the end of the trial (7% of those randomised, fig 1). Sociodemographic and clinical data were collected for all randomised participants. The proportion returning questionnaire data was lowest in the middle of the trial: for some participants we could administer the questionnaires only at the start and end of the trial because of lengthy gaps between appointments. Eight patients from the specialist breast unit and 17 from the district general hospital did not return questionnaires at any time point. Among the 349 other patients, participation at the end of the trial was not associated with any of the variables in tables 1 and 2 at 10% significance using logistic regression models, apart from time from diagnosis (likelihood ratio  $\chi^2=3.28$ ,  $df=1$ ,  $P=0.070$ ) and time from first visit after treatment ( $\chi^2=2.90$ ,  $df=1$ ,  $P=0.089$ ). Those participating at the end of the trial had a shorter time from diagnosis at the start of the trial (median 18 months) than those who had dropped out after participating at least once (median 25.5 months, Mann-Whitney  $U=5902.5$ ,  $P=0.019$ ), and a shorter time from post-treatment visit at the start of the trial (15 v 25 months,  $U=6046.0$ ,  $P=0.033$ ).

**Table 3** | State-trait anxiety inventory (STAI) findings by study group adjusted for site and follow-up status (analysis by intention to treat or adjusted for treatment received)

Group	No of women	Estimated marginal mean (95% CI)	$F_{df}$ , P value	95% CI for difference*
<b>Intention to treat</b>				
Start of trial (STAI Y1):				
Hospital	158	36.45 (34.42 to 38.49)	0.05 <sub>1,320</sub> , 0.821	-2.88 to 2.28
Telephone	167	36.16 (34.12 to 38.20)		
Start of trial (STAI Y2):				
Hospital	156	39.20 (37.19 to 41.21)	1.84 <sub>1,313</sub> , 0.176	-4.32 to 0.79
Telephone	162	37.43 (35.41 to 39.46)		
Middle of trial:				
Hospital	108	34.95 (32.20 to 37.70)	1.11 <sub>1,228</sub> , 0.293	-4.64 to 1.41
Telephone	125	33.33 (30.59 to 36.08)		
End of trial:				
Hospital	132	35.65 (33.36 to 37.95)	0.26 <sub>1,271</sub> , 0.608	-3.61 to 2.18
Telephone	144	34.91 (32.65 to 37.16)		
Start, middle, and end of trial†:				
Hospital	398	36.13 (34.33 to 37.93)	0.24 <sub>1,326.2</sub> , 0.623	-2.82 to 1.69
Telephone	436	35.57 (33.78 to 37.36)		
<b>Adjusted for treatment received</b>				
Middle of trial:				
Hospital	111	35.07 (32.18 to 37.96)	1.13 <sub>1,227</sub> , 0.290	-5.42 to 1.63
Telephone	122	33.18 (30.27 to 36.09)		
End of trial:				
Hospital	136	35.67 (33.20 to 38.14)	0.22 <sub>1,270</sub> , 0.637	-4.27 to 2.62
Telephone	140	34.85 (32.37 to 37.34)		
Start, middle, and end of trial†:				
Hospital	413	36.15 (34.23 to 38.08)	0.21 <sub>1,325.8</sub> , 0.648	-3.33 to 2.07
Telephone	421	35.52 (33.56 to 37.49)		

\*Telephone group minus hospital group (negative differences imply less anxiety in telephone group).

†Linear mixed model. Mean difference from start to middle of trial -0.66, paired  $t=-1.04$ ,  $df=225$ ,  $P=0.301$ , 95% CI -1.90 to 0.59. Mean difference from start to end of trial -0.24, paired  $t=-0.36$ ,  $df=266$ ,  $P=0.718$ , 95% CI -1.52 to 1.05.

### Psychological morbidity

Differences between groups in state-trait score were not significant at the start, middle, or end of the trial under intention to treat or adjusted treatment received analyses, although means were consistently lower for the telephone group (table 3). Mean score did not significantly improve during the trial, the mean reduction from the start to the end of the trial being 0.33.

Figure 2 shows 95% confidence intervals for the differences between mean values (hospital v telephone) of state-trait scores for intention to treat and adjusted treatment received analyses with linear mixed models. Positive differences in means indicate more anxiety in the telephone group, while negative differences indicate more anxiety in the hospital group. Both confidence intervals were within the pre-stated equivalence region of -3.5 to 3.5, indicating equivalence.<sup>22</sup>

We compared groups for other variables under intention to treat analysis. GHQ-12 scores were highly skewed, with over 49% of scores within a group being 0 at each time point. Differences between randomised groups at the start, middle, or end of the trial were not significant, nor were differences between time points. Although the percentage of cases (scores  $\geq 4$ ) was consistently higher in the hospital group at the start, middle, and end of the trial, differences between the groups at each time point were not significant. Initially, 71/328 (22%) were GHQ-12 cases compared with 47/

**Table 4** | Information needs at start, middle, and end of trial by randomised group. Figures are numbers (percentages) of women

Time in trial	Hospital	Telephone
Spread of disease:		
Start	62/170 (37)	70/175 (40)
Middle	19/121 (16)	24/136 (21)
End	25/145 (17)	32/153 (21)
Treatment and side effects:		
Start	64/168 (38)	71/174 (41)
Middle	22/121 (18)	22/137 (16)
End	24/144 (17)	30/153 (18)
Genetic risk:		
Start	74/168 (44)	76/174 (44)
Middle	34/121 (28)	43/137 (31)
End	43/144 (30)	49/152 (32)
Sexual attractiveness:		
Start	30/170 (18)	37/174 (21)
Middle	18/121 (15)	9/137 (7)
End	15/144 (10)	20/153 (13)
Self care:		
Start	55/170 (32)	58/175 (33)
Middle	18/121 (22)	26/136 (19)
End	25/144 (17)	23/153 (15)
Family concerns:		
Start	48/170 (28)	40/175 (23)
Middle	21/121 (17)	19/138 (14)
End	20/145 (14)	16/153 (11)

281 (17%) at the end of the trial. Change for the 266 with data at both time points was not significant.

### Information needs

Participants clearly indicated their specific information needs (table 4). Initially, the highest need related to information about genetic risk, the lowest for information on sexual attractiveness. Within both randomised groups, information needs reduced over time for all items. There was little difference between the groups in information needs, apart from information on sexual attractiveness in the middle of the trial. The need for information on genetic risk remained the highest at the end of the trial, with 92 of the 296 respondents to the question (31%) still requiring information.

### Participants' satisfaction

There were no significant differences between randomised groups initially regarding satisfaction with information received (table 5). The telephone group showed significantly more satisfaction at the middle and end of the trial ( $P < 0.001$ ). Participants were asked if they had thought that the appointment had been helpful in dealing with their concerns. There was no difference between groups initially but at the middle and end of the trial, responses were significantly more positive in the telephone group, with a higher percentage reporting "very helpful" and few with negative responses (table 6).

Table 7 shows the numbers of contacts with health professionals or administrative staff between appointments for both groups; contacts were relatively few but were primarily with breast care nurses, general practitioners, and lymphoedema nurses. There were no significant differences between groups in terms of contact at any time point. Women who had contact with their general practitioners, however, tended to have been diagnosed more recently (median 10 months compared with 20 months for those with no reported contact, Mann-Whitney  $U = 2044.0$ ,  $P = 0.004$ ). Women who had contact with breast care nurses between appointments also tended to have been diagnosed more recently (median 6.5 months compared with 20 months for those with no reported contact,  $U = 1488.5$ ,  $P = 0.001$ ). Contact with lymphoedema nurses was not associated with time from diagnosis ( $U = 1895.5$ ,  $P = 0.244$ ). There were few contacts with hospital doctors and administrative staff between visits, and numbers were too small for additional analysis.

### Clinical investigations ordered

There were no differences between groups in whether clinical investigations were ordered for participants as a result of appointments at the start (hospital 29% *v* telephone 24%,  $\chi^2 = 1.10$ ,  $df = 1$ ,  $P = 0.294$ ), middle (36% *v* 34%,  $\chi^2 = 0.08$ ,  $df = 1$ ,  $P = 0.772$ ), or end of the trial (40% *v* 43%,  $\chi^2 = 0.32$ ,  $df = 1$ ,  $P = 0.574$ ). In most cases, investigations comprised routine mammograms. Other investigations mentioned by both groups at all stages of the study included 18 non-routine mammograms, 13 blood tests, nine chest x ray investigations, nine bone scans, six fine needle aspirations/biopsies, and one magnetic resonance imaging scan.

### Time to detection of recurrence

Only 17 participants (5%) had a confirmed recurrence of cancer during the trial: six in the hospital group and 11 in the telephone group (table 8). The difference between randomised groups was not significant ( $\chi^2 = 1.33$ ,  $df = 1$ ,  $P = 0.250$ ). The median time to confirmation was 60.5 days (range 37-131 days) in the hospital group and 39.0 days (10-152 days) in the telephone group (Mann-Whitney  $U = 21.0$ ,  $P = 0.228$ , for difference).

## DISCUSSION

### Positive benefit

Telephone follow-up by specialist breast care nurses has positive benefits for women with breast cancer. Our study was specifically designed to meet the information needs of patients. It was encouraging that nearly 60% agreed to participate, given that patients find hospital visits reassuring<sup>10</sup> and they were asked to forego clinical examinations. Routine mammograms took place irrespective of group allocation and this might have had an impact on patients' preferences for participation, particularly for those patients with lesions detected at screening who had no symptoms on initial presentation.

**Table 5** | Satisfaction with information received by randomised group (intention to treat analysis). Figures are numbers (percentages) of women

Level of satisfaction	Hospital	Telephone	Between groups comparison P value	
			All categories*	First 4 categories
<b>Start of trial</b>				
Very satisfied	79 (46)	78 (45)	0.671	$\chi^2_{\text{trend}}=0.31, 0.576$
Satisfied	67 (39)	71 (41)		
Not very satisfied	4 (2)	8 (5)		
Very unsatisfied	1 (1)	0 (0.0)		
Did not receive information	10 (6)	6 (4)		
Did not need information	10 (6)	10 (6)		
Overall	171 (100)	173 (100)		
<b>Middle of trial</b>				
Very satisfied	57 (49)	110 (80)	<0.001	$\chi^2_{\text{trend}}=19.07, <0.001$
Satisfied	41 (35)	22 (16)		
Not very satisfied	7 (8)	3 (2)		
Very unsatisfied	3 (3)	1 (1)		
Did not receive information	4 (3)	0 (0)		
Did not need information	4 (3)	1 (1)		
Overall	116 (100)	137 (100)		
<b>End of trial</b>				
Very satisfied	78 (55)	121 (80)	<0.001	$\chi^2_{\text{trend}}=14.33, <0.001$
Satisfied	47 (33)	25 (16)		
Not very satisfied	8 (6)	3 (2)		
Very unsatisfied	1 (1)	1 (1)		
Did not receive information	7 (5)	0 (0)		
Did not need information	1 (1)	2 (1)		
Overall	142 (100)	152 (100)		

\*Fisher's exact test.

Our study was not about finding the most useful or sensitive tests and investigations for detecting recurrence. We focused on the consultation between participants and clinicians, aiming to provide participants with the information and support they needed, when they needed it, to cope with the diagnosis of cancer. The telephone intervention provided a service to participants that met their needs, with no evidence of physical or psychological disadvantage. This trial was primarily designed to evaluate whether there was equivalence, in terms of psychological morbidity, between hospital and telephone follow-up, and this was shown under the more conservative approach adjusted for treatment received.

Scores on the state-trait anxiety inventory did not significantly improve in either group during the trial, and this might indicate that, although this measure has been used successfully with patients with breast cancer in previous studies, it was not sufficiently sensitive to capture changes over time in this study. Findings for the anxiety inventory and GHQ-12 indicated that there were no differences between scores for study groups at any point in the trial. Participants who received telephone follow-up were not more anxious as a result of foregoing hospital contact and clinical examinations.

Those in the telephone group reported greater satisfaction with the information received and reported appointments as more helpful in meeting their needs. The telephone intervention was specifically designed to provide information and hence met its objectives.

The percentage of participants requiring information on specific needs at the end of the trial ranged from 10% to 32%. Considering that participants were a median of 20 months from diagnosis at the point of recruitment and remained in the study for a mean of 24 months,

**Table 6** | Helpfulness in dealing with concerns at appointment by randomised group (intention to treat analysis). Figures are numbers (percentages) of women

Level of satisfaction	Hospital	Telephone	$\chi^2_{\text{trend}}$ , P value
<b>Start of trial</b>			
Very helpful	44 (48)	44 (52)	0.69, 0.405
Helpful	36 (39)	31 (37)	
Not very helpful	10 (11)	9 (11)	
Very unhelpful	2 (2)	0 (0)	
Overall	92 (100)	84 (100)	-
<b>Middle of trial</b>			
Very helpful	28 (44)	80 (88)	28.27, <0.001
Helpful	26 (41)	9 (10)	
Not very helpful	9 (14)	1 (1)	
Very unhelpful	1 (2)	1 (1)	
Overall	64 (100)	91 (100)	-
<b>End of trial</b>			
Very helpful	42 (63)	84 (84)	10.35, 0.001
Helpful	15 (22)	13 (13)	
Not very helpful	10 (15)	2 (2)	
Very unhelpful	0 (0)	1 (1)	
Overall	67 (100)	100 (100)	-

**Table 7** | Reported number of contacts with health professionals between appointments by randomisation group

	Hospital	Telephone
<b>Breast care nurses</b>		
Baseline	14	26
Middle of trial	13	18
End of trial	15	14
<b>General practitioners</b>		
Baseline	14	13
Middle of trial	12	9
End of trial	14	13
<b>Lymphoedema nurses</b>		
Baseline	10	13
Middle of trial	10	7
End of trial	14	19
<b>Hospital doctor</b>		
Baseline	4	3
Middle of trial	3	6
End of trial	6	3
<b>Hospital administrative/secretarial staff</b>		
Baseline	4	7
Middle of trial	3	4
End of trial	5	4
<b>Community nurses</b>		
Baseline	3	1
Middle of trial	0	1
End of trial	3	1

**Table 8** | Characteristics of patients with confirmed recurrence of cancer by randomised group (intention to treat analysis)

	Hospital	Telephone	P value*
<b>Study site:</b>			
District general hospital	5	10	0.999
Specialist breast unit	1	1	
<b>Type of recurrence:</b>			
Local	4	4	0.335
Distant metastases	2	7	
<b>Patient died (related to breast cancer):</b>			
Yes	2	6	0.620
No	4	5	
<b>Presentation:</b>			
Contacted general practitioner	3	6	0.891
Phoned breast care nurse	1	1	
Presented symptoms to breast care nurse during pre-arranged appointment	0	2	
Routine mammogram	2	2	
<b>Routine/interval visit:</b>			
Routine appointment, patient symptomatic	0	2	0.787
Interval visit, patient with symptoms	4	7	
Interval presentation, routine mammogram	2	2	
Totals	6	11	-

\*Fisher's exact test.

many patients retained a need for information long after completion of treatment. Breast care nurses in the telephone arm of the study received training in administering the structured intervention, designed to meet information needs. Provision of training in meeting information needs for all health professionals involved in providing follow-up care for women treated for breast cancer could be beneficial.

There were no differences in terms of investigations ordered between groups. A lack of visual cues did not result in more tests being ordered. These findings, however, are based on participants' retrospective recall of investigations ordered; response might have been subject to inaccuracies. As part of this study we collected more detailed data on resource use to inform an economic evaluation that will present a more detailed and accurate picture of tests and investigation. Participants in the telephone group were no more likely to consult with other health professionals between visits than those in the hospital group and so were not using additional healthcare resources. Women who initiated contact with breast care nurses and general practitioners between appointments tended to be less than a year from diagnosis; negotiating frequency of appointments with patients might be appropriate to meet differing needs at different time points.

#### Strengths and limitations

The study provides limited information about time to detection of recurrent disease as the participants had a low to moderate risk of recurrence; only 17 recurrences were detected. These were mostly interval events; no recurrences were detected in patients without symptoms at routine appointments. There were no differences in time to detection of recurrence between the groups. Despite appointments taking place over the telephone, there were no undue delays in identifying potential clinical problems and instigating appropriate referral processes. Given these positive findings, a similar approach could be considered for patients in all risk categories as follow-up does not usually have a different format for high risk groups. At this stage, however, our findings can be applied with confidence only to women at low risk of recurrence; we did not have a large enough sample to provide robust data on time to detection of recurrence, and larger studies would be needed across risk groups to determine if telephone follow-up could be effective for women at high risk of recurrence. More work would also identify the specific needs of patients at high risk of recurrence with subsequent adaptation of the intervention instrument.

The telephone intervention presented a structured approach to follow-up, which could have several advantages. Information could be repeated at each appointment; patients might not recall information previously given or might find repetitive information reassuring. The structure allows patients to choose what information is important to them and when it is important. Administration of the telephone intervention requires training as high levels of skill and



## WHAT IS ALREADY KNOWN ON THIS TOPIC

UK national guidelines recommend that routine long term follow-up after treatment for breast cancer is not effective at prolonging survival

Brief consultations aimed at detection of recurrence do not provide opportunities for discussion of information and psychosocial needs

## WHAT THIS STUDY ADDS

Participants in a telephone follow-up group were no more anxious as a result of foregoing clinical examinations and face-to-face contact

Telephone follow-up was associated with high levels of satisfaction in patients

Telephone follow-up might decrease the burden on busy hospital clinics

knowledge are essential, and breast care nurses are ideally and uniquely placed for this role. This study was conducted within the confines of clinical reality, given all the usual challenges of staff shortages and sickness; no additional nursing services were provided. Specialist nurses would probably not have the capacity to provide telephone follow-up to all women treated for breast cancer, even if eligibility was limited to those at low risk of recurrence. Telephone follow-up might reduce the burden on busy hospital clinics, but other approaches might also be effective. Alternating telephone and hospital follow-up, according to patients' preferences, could be a suitable approach, while hospital follow-up might continue to be preferred by those who do not feel comfortable discussing their concerns over the telephone.

Telephone follow-up is convenient, especially in rural areas where patients might have to travel long distances for hospital appointments and for those with limited mobility. Patients might also feel more comfortable in their own homes, where they cannot see the "busyness" of hospital clinics, and might take the opportunity to be more proactive in seeking answers to their questions. Telephone follow-up might have broader applicability to other groups of patients. There are nurse specialists for many diseases and their skills could be harnessed to provide a quality service while reducing the burden on busy hospital clinics.

We thank the following collaborators whose continued support was vital to the successful completion of the study: A Baildam, L Barr, G Byrne, V Chadwick, F Danwata, S Foster, P Kiriparan, M Noblet, F O'Regan, S Rajan, L Thomson, C Turner. Lesley Degner's seminal work on information needs of cancer patients has provided continued inspiration. We also thank all the patients who participated in this study.

**Contributors:** KB was principal investigator and together with MC, GD, AH, KL, SS, MT, and DTR contributed to the study concept and design. KB, MT, and SW took main responsibility for the acquisition of data; MC and GD were responsible for data analysis and all authors were responsible for interpretation of findings. KB, MC, and KL drafted the manuscript and all authors contributed to the critical revision of the manuscript. KB, GD, AH, KL, and SS obtained funding. KB is the guarantor.

**Funding:** The trial was funded by the Medical Research Council (UK) and a project grant from Rosemere Cancer Foundation (UK). The funding agencies had no role in the design and conduct of the study, in the collection, analysis and interpretation of the data, or in the preparation, review or approval of the manuscript

**Competing interests:** None declared.

**Ethical approval:** This study was approved by the central office of research ethics committees in England and research and development

departments at both study sites. All participants gave both verbal and written consent to take part in the study.

**Provenance and peer review:** Not commissioned; externally peer reviewed.

- Rojas MP, Telaro E, Russo A, Moschetti I, Coe L, Fossati R, et al. Follow up strategies for women treated for early breast cancer. *Cochrane Database Syst Rev* 2005;(1):CD001768.
- te Boekhorst DS, Peer NG, van der Sluis RF, Wobbes T, Ruers TJ. Periodic follow-up after breast cancer and the effect on survival. *Eur J Surg* 2001;167:490-6.
- Ashkanani F, Sarkar T, Needham G, Coldwells A, Ah-See AK, Gilbert FJ et al. What is achieved by mammographic surveillance after breast conservation treatment for breast cancer? *Am J Surg* 2001;182:207-10.
- Chum M, Kelly V. Outpatient follow-up after treatment for early breast cancer: updated results after 5 years. *Clin Oncol* 2001;13:187-94.
- Dewar J, Kerr G. Value of routine follow up of women treated for early carcinoma of the breast. *BMJ* 1985;291:1464-67.
- Donnelly J, Mack P, Donaldson LA. Follow-up of breast cancer: time for a new approach? *Int J Clin Pract* 2001;55:431-3.
- Grunfeld E, Mant D, Yudkin P, Adewuyi-Dalton R, Cole D, Stewart J, et al. Routine follow up of breast cancer in primary care: randomised trial. *BMJ* 1996;313:665-9.
- Snee M. Routine follow-up of breast cancer patients. *Clin Oncol* 1994;6:154-6.
- National Institute for Health and Clinical Excellence. *Improving outcomes in breast cancer*. London: NICE, 2002.
- Beaver K, Luker KA. Follow-up in breast clinics: reassuring for patients rather than detecting recurrence. *Psychooncology* 2005;14:94-101.
- Grunfeld E, Many D, Vessey MP, Fitzpatrick. Specialist and general practice views on routine follow-up of breast cancer patients in general practice. *Fam Pract* 1995;12:60-5.
- Degner LF, Kristjanson LJ, Bowman D, Sloan JA, Carriere KC, O'Neil J et al. Information needs and decisional preferences in women with breast cancer. *JAMA* 1997;277:1485-92.
- Luker KA, Beaver K, Leinster SJ, Owens RG. The information needs of women with breast cancer: a follow up study. *J Adv Nurs* 1996;23:487-95.
- McArdle J, George WD, McArdle CS, Smith DC, Moodie AR, Hughson AV, et al. Psychological support for patients undergoing breast cancer surgery: a randomised study. *BMJ* 1996;312:813-6.
- Baildam A, Keeling F, Noblet M, Thomson L, Bundred N, Hopwood P. Nurse led follow-up clinics for women treated for primary breast cancer: a randomised controlled trial. *Eur J Cancer* 2002;38(suppl 3):136.
- Gotay C, Bottomley A. Providing psycho-social support by telephone: what is its potential in cancer patients? *Eur J Cancer Care* 1998;7:225-31.
- Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW. Reporting of noninferiority and equivalence randomised trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.
- Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham prognostic index in primary breast cancer. *Breast Cancer Res Treat* 1992;22:207-19.
- Beaver K, Twomey M, Witham G, Foy S, Luker K. Meeting the information needs of women with breast cancer: piloting a nurse led intervention. *Eur J Oncol Nurs* 2006;10:378-90.
- Spielberger CD. *State-trait anxiety inventory*. Mountain View, CA: Consulting Psychologists Press, 1983.
- Goldberg DP. *Manual of the general health questionnaire*. Windsor: NFER Nelson, 1978.
- Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;313:36-9.
- Armitage P, Berry G, Matthews JNS. *Statistical methods in medical research*. 4th ed. London: Blackwell Science, 2001.
- White IR. Uses and limitations of randomization-based efficacy estimators. *Stat Meth Med Res* 2005;14:327-47.
- Nagelkerke N, Fidler V, Bemsens R, Borgdorff M. Estimating treatment effects in randomised clinical trials in the presence of non-compliance. *Stat Med* 2000;19:1849-64.
- Hausman JA. Specification tests in econometrics. *Econometrica* 1978;46:1251-71.
- Dunn G, Bentall R. Modelling treatment-effect heterogeneity in randomised controlled trials of complex interventions (psychological treatments). *Stat Med* 2007;26:4719-45.
- Norusis MJ. *SPSS 15 advanced statistical procedures companion*. Upper Saddle River, NJ: Prentice Hall, 2007.
- Kitchener H, Dunn G, Lawton V, Reid F, Nelson L, Smith A. Laparoscopic versus open colposuspension—results of a prospective randomised controlled trial. *Br J Obstet Gynaecol* 2006;113:1007-13.
- Office for National Statistics. *Health: breast cancer*. www.statistics.gov.uk/cci/nugget.asp?id=575.

Accepted: 15 October 2008