carcinoma. Final report of Head and Neck Contracts Program. *Cancer* 1987, **60**, 301–311.

7. Schuller DE, Wilson H, Hodgson S, Mattox D. Preoperative reductive chemotherapy for stage III or IV operable epidermal carcinoma of the oral cavity, oropharynx, hypopharynx or larynx: phase III. A southwest oncology group study (abstr.). *Proc Int Conf Head and Neck Cancer* 1984, **1**, 48.

8. Toohill RJ, Anderson T, Byhardt RW, *et al.* Cisplatin and fluorouracil as neoadjuvant therapy in head and neck cancer. A preliminary report. *Arch Otolaryngol Head Neck* 1987, **113**, 758–761.

9. Kish J, Ensley JF, Jacob S *et al.* A randomised trial of cisplatin (CACP) + 5 fluorouracil infusion and CACP + 5 FU bolus for recurrent and advanced squamous cell carcinoma of the head and neck. *Cancer* 1983, **56**, 2740–2741.

10. Domenge C, Marandas P, Leridant AM, Richard JM. Importance du choix de la dose dans le traitement des carcinomes épidermoides de la cavité buccale et de l'oropharynx. Forum de Cancérologie (abstr.) *Bull Cancer* 1986, **73**, 470.

11. Ervin TJ, Clark JT, Weischselbaum PR, *et al.* An analyse of induction and adjuvant chemotherapy in the multidisciplinary treatment of squamous-cell carcinoma of the head and neck. *J Clin Oncol* 1987, **5**, 10–20.

Development and Validation of an Instrument to Measure Satisfaction of Participants at Breast Screening Programmes

Jill Cockburn, David Hill, Les Irwig, Trudy De Luise, Deborah Turnbull and Penny Schofield

A reliable and valid questionnaire has been developed to measure the satisfaction of participants with service offered at mammography screening programmes. The questionnaire measures five specific aspects: convenience and accessibility, staffs' interpersonal skills, information transfer between staff and client, physical surroundings and perceived technical competence of staff. A general satisfaction dimension was also included. Systematic procedures were followed to ensure that the initial pool of items met the criteria for satisfactory content validity. These procedures included extensive literature review and interviews with participants and service providers. Discriminant validity was assessed by a modified Q-sort procedure, where eight expert judges sorted items into relevant dimensions. The sample for other validity and reliability testing consisted of 584 women who were participants at a breast X-ray programme in Melbourne, Australia. Concurrent validity was demonstrated by considering the correlation of the sum of the subscale scores for each respondent with their score on the general subscale ($r = 0.76$; $P < 0.001$). Multiple regression was used to provide further evidence for the discriminant validity of the proposed subscales and support for the multidimensional conceptualism of satisfaction. Scores on the general satisfaction subscale were used as an outcome variable and other subscale scores were predictor variables. All subscale scores significantly contributed to the prediction of satisfaction, over and above that of other subscales ($R^2 = 0.59$). This indicates that these subscales are measuring distinct dimensions of satisfaction. Cronbach's alpha of each subscale was over 0.50, indicating that the subscales are reliable. The instrument is a potentially useful tool for assessing the quality of care at mammographic screening services and could be used routinely by such services to monitor satisfaction.
Eur J Cancer, Vol. 27, No. 7, pp. 827–831, 1991

INTRODUCTION

BREAST CANCER is one of the most common causes of death in many western countries [1]. Mass mammography screening has the potential to reduce the death rate from the disease. A

reduction in mortality of about one third in populations of women over about 50 offered screening has been demonstrated [2, 3]. The United Kingdom has therefore introduced population screening for women in the appropriate age groups. Among the key elements of the quality assurance goals of the National Health Service breast screening programme is "minimising dissatisfaction and striving to ensure that women find screening a positive and healthy experience" [4].

There are a number of reasons why the measurement of satisfaction is essential. First, ongoing evaluation of health services is necessary to ensure that participants are being offered the best possible service. The degree of satisfaction expressed

Correspondence to J. Cockburn.
J. Cockburn, D. Hill, T. De Luise and P. Schofield are at the Centre for Behavioural Research in Cancer, Anti-Cancer Council of Victoria, 1 Rathdowne St, Carlton Sth, Victoria 3053; and L. Irwig and D. Turnbull are at the Department of Public Health, University of Sydney, NSW, Australia.
Revised 14 Feb. 1991; accepted 9 Apr. 1991.

by participants is one reflection of the standard and quality of the service given. This reason alone is sufficient justification for making the monitoring of participant satisfaction an integral part of the delivery of a mammography screening program.

However, the degree of satisfaction reported by participants has more far-reaching consequences in that patterns of attendance for both initial and repeat screening may be contingent on these perceptions. A number of factors have been shown to be related to non-attendance for screening mammography. These include personal characteristics of the woman, such as perceptions of susceptibility to breast cancer, the fear of outcome and perceptions of efficacy of the procedure [5, 6], as well as structural barriers, such as distance from a centre [7] or the lack of invitation because of incomplete population registers [8]. However, there can be no denying that a woman's own evaluation of the service offered at a screening centre, in terms of the way she is dealt with by the staff, the amount of information given to her, and the accessibility of the service, will be a potent determinant of her decisions about attendance for rescreening. Research on patient satisfaction with medical care suggests that patient dissatisfaction leads to delay in seeking medical treatment [9] and non-adherence to treatment instructions [10]. Satisfaction may be even more important in the case of screening mammography where participants are not motivated by ill-health to comply and where enduring benefit to the participant is dependent upon remaining in the screening program over many years.

Moreover, women's evaluations of the service may also have more widespread implications for attendance at initial screening. An important means by which women hear about the service is through friends and neighbours attending. Positive opinions of these social referents are associated with the intention to have a mammogram [5]. It is therefore imperative that "word of mouth" about the pogramme is complimentary; criticism may dissuade others from attending. The regular monitoring of satisfaction allows for any problems to be detected and dealt with by appropriate changes in service delivery.

To date there has been little research on the effect of satisfaction on attendance for initial and repeat screens. One reason for this may be the current lack of adequate means of measuring satisfaction. Given its importance, it is essential that the measurement of participant satisfaction is methodologically rigorous, otherwise spurious results will be obtained possibly leading to inaccurate decisions about changes in service delivery. There are a number of issues involved in the development of an adequate instrument for measuring satisfaction [11]. First, satisfaction should be treated as a multidimensional concept. Studies of satisfaction with medical care have shown that different aspects of health care delivery, e.g. accessibility of services, provider conduct and physical environment, can shape participant satisfaction [12]. Moreover, different levels of satisfaction may be reported for different dimensions of care [13].

Secondly, measurements should be reliable. Reliability refers to the extent to which measurements are repeatable. To enhance reliability, scales with multiple items to measure each of the different dimensions of satisfaction should be used. Single-item scores lack precision and are substantially influenced by other methodological considerations in addition to the construct being measured [14]. For example, single items are influenced by acquiescence response set—the tendency of respondents to agree with statements regardless of their content. Acquiescence biases results since it causes significantly depressed scores with negatively worded items. Balanced scales with approximately equal

numbers of negatively and positively worded items within scales have been recommended as a counter to this kind of systematic bias [14].

Finally, instruments should be valid, that is, they should measure what they are purported to measure. There are several types of validity criteria including predictive, concurrent, content, and discriminant validity [15, 16]. Predictive validity will be addressed in future studies relating satisfaction scores with subsequent behaviour such as reattendance, and is not further considered in this paper. Where there are no published satisfaction scales which would be appropriate for comparison, assessment of concurrent validity is limited to comparing alternate methods of questioning and/or correlating the sum of subscales with a general scale measured concurrently. Content validity refers to the extent which items fairly represent the domain of the phenomenon under study, and is assessed by comparing items logically to the content being measured. Discriminant validity involves determining whether respondents can discriminate between separate dimensions. If evidence of such discrimination is found then different dimensions should be measured and reported separately. One way to do this is to assess whether respondents are evaluating each aspect of the service independently of other aspects [17].

The aim of this study was to develop a questionnaire using accepted psychometric procedures to measure the satisfaction of participants at mammographic screening programmes.

METHODS AND RESULTS

Development of item pool

Systematic procedures were followed in developing a pool of items to meet criteria for satisfactory content validity. Based on a review of the literature on patients' satisfaction with medical care [12, 13, 17–19] and interviews with staff members including the project director, radiologist, radiographers and medical staff at a mammography screening service, five specific and one general dimension of satisfaction were identified. Specific dimensions were: convenience and accessibility, staffs' interpersonal skills, information transfer between staff and client, physical surroundings and perceived technical competence of staff. Content of the initial pool of items to measure these dimensions came from a review of published scales which measure satisfaction with medical care and interviews with staff and clients at both a private breast clinic and a mammographic screening programme. The 10 staff members included the project director, radiologists, radiographers and medical staff. Approximately 50 women were interviewed over three sessions at the private breast clinic, and 15 women interviewed over two sessions at the mammographic screening programme. Items derived from these sources were edited to reduce repetitions and ambiguities, leaving 37 items in the pilot form.

Testing the pilot form

Items in the pilot instrument were structured in the form of a complete statement, e.g. "the staff were very friendly". A five-point Likert response format ranging from strongly agree to strongly disagree was used for all items. Items from each dimension were randomly distributed throughout the questionnaire and responses were marked directly onto the questionnaire.

15 women who were consecutive attenders at a mobile mammographic screening service in Sydney, Australia took part in testing the pilot form. A research assistant sat with these women as they completed the questionnaire, asking women for comments about any specific problems with particular questions.

Table 1. Contents of the breast screening satisfaction scale

Items (in order presented)*	Dimension†
1. I felt free to ask the staff questions I wanted to ask	SIS
2. The person was too rough when taking the X-ray (N)	PTC
3. I felt that I had to wait far too long (N)	CA
4. I was very satisfied with the care I received at the service	GS
5. I feel confident that the X-ray was taken properly	PTC
6. The staff seemed to hurry me through too quickly (N)	SIS
7. I was told all I wanted to know about what was done at the service	IT
8. The staff did everything they could to make me feel comfortable about having the X-ray	SIS
9. The service is in a place which is easy for me to get to	CA
10. The staff had good manners	SIS
11. I feel I had a chance to speak freely to the staff	IT
12. The staff used words that were hard to understand (N)	IT
13. The temperature in the place was uncomfortable (N)	PS
14. I found it hard to find a convenient time to come to the service (N)	CA
15. The service I got was much worse than I expected (N)	GS
16. I am sure that the reception staff knew what they were doing	PTC
17. The hours which the service is open suit me	CA
18. I would strongly recommend the service to my friends	GS
19. I had enough privacy while the X-ray was being done	PS
20. The staff told me all I wanted to know about when I would get results back	IT
21. The changing rooms were too small (N)	PS
22. I could not find any faults in the service I received	GS
23. I found the waiting room very pleasant	PS
24. I think the service could be run much more efficiently (N)	GS
25. The person who took the X-ray seemed to know what s/he was doing	PTC
26. I had enough privacy while getting changed	PS

*Response options and scale values (which are reverse for negatively worded items indicated by 'N') were as follows: strongly agree (1), agree (2), not sure (3), disagree (4) and strongly disagree (5).
†Convenience and accessibility (CA); staffs' interpersonal skills (SIS); information transfer between staff and client (IT); physical surroundings (PS); perceived technical competence of staff (PTC); and general satisfaction (GS).

Some items were eliminated and the wording of others modified on the basis of these results. A further sample of 18 women at the Sydney mammography service were given the modified questionnaire to complete after their mammogram. The research assistant again gathered comments about any difficulties with

Table 2. Cronbach's alpha for subscales of the BSSS

Subscales	Alpha
Convenience and accessibility	0.53
Staffs' interpersonal skills	0.70
Information transfer	0.58
Physical surroundings	0.68
Perceived technical competence	0.56
General	0.67

completing the questionnaire. Minor modifications to wording were made at this stage.

As one part of our assessment of the scales' discriminant validity, a modified Q sort procedure [16] was used to confirm that postulated dimensions within the scale could be distinguished on the basis of item content. Each item was written on a card and the cards were submitted independently to eight judges. These judges included clinical psychologists and other research staff. They were asked to sort the items into the domain which they seemed to represent. Items misclassified by more than 20% of the judges were excluded from the questionnaire at this stage. One item, "the service could be run more efficiently", was originally intended to measure the technical competence dimension. However, all judges classified it as a general satisfaction item. It was therefore retained in the questionnaire as a measure of the general satisfaction dimension.

Table 1 lists the 26 items in the final questionnaire, which we call the breast screening satisfaction scale (BSSS). There were four items in each of the convenience and accessibility, interpersonal skills, information transfer and perceived technical competence subscales. The general satisfaction and physical surrounding subscales each contained five items. At least one of the items within each subscale were negatively worded to control for acquiescence bias.

Testing the final form

Respondents in this study were participants at a mammography screening service in Melbourne, Australia. Approximately 100 consecutive participants at six separate time periods, 3 months apart, were approached and asked to participate. Respondents were handed the BSSS as they were leaving and were asked to take it home and complete it. A reminder phone call was made 48 hours after the clinic visit. Of the 608 women approached, 584 (96%) returned the questionnaire. As well as the standard 26 Likert scale items, the questionnaire also included open-ended questions about each dimension, for women to record any problems they found with these particular aspects of service delivery.

Reliability of the scale was assessed by calculating Cronbach's coefficient alpha [20] and the results of this analysis are represented in Table 2. The Cronbach's alpha values are all above the accepted limit (0.50) which indicates internal consistency [21].

Concurrent validity was assessed by correlating the sum of the subscale scores for each respondent with their score on the general scale. All scores were transformed into z-scores for all analyses. The correlation between the sum of the subscale scores and the scores on the general scale was $r = 0.76; P < 0.0001$.

Clearly, since measures of components of satisfaction form part of the whole, each dimension must be shown to relate with

Table 3. *Association of satisfaction dimensions with measure of general satisfaction*

Subscale	Bivariate association		Association adjusted for all other subscale items	
	Pearson's r	P	Partial r	P
Competence	0.68	<0.0001	0.31	<0.0001
Physical surroundings	0.63	<0.0001	0.18	<0.0001
Interpersonal skills	0.66	<0.0001	0.16	0.0002
Information transfer	0.63	<0.0001	0.18	<0.0001
Convenience	0.48	<0.0001	0.08	0.06

the general measures. Table 3 presents the Pearson correlations of each subscale with the general measure. It can be seen that correlations with general satisfaction range from 0.65 for the competence subscale to 0.48 for the convenience subscale. All correlations were significant beyond the 0.0001 level.

A multiple regression was used to examine the discriminant validity of the proposed subscales and possible support for the multidimensional conceptualisation of satisfaction. Scores on the general satisfaction subscale were used as the outcome variable and other subscale scores were predictor variables [17]. The purpose of such an analysis was to test that hypothesised dimensions contributed independently to explain the variance of the general factor. The results of the regression are also shown in Table 3.

The final equation accounted for 59% of the measured variance of the general satisfaction scale. The partial correlations ranged from 0.31 for the competence subscale to 0.08 for the convenience subscale. For all subscales, except convenience, the partial correlation with the general subscale was significant at $P < 0.0002$. The level of significance of the partial correlation between the convenience subscale and the general satisfaction subscale was $P = 0.06$. As this is close to conventional levels of significance it seems appropriate to include it in the final equation [22]. Therefore, each subscale significantly contributed independently to the prediction of general satisfaction. This indicates that these subscales are measuring distinct dimensions of satisfaction.

DISCUSSION

We have described the development of an instrument that is an internally consistent and empirically valid measure of attenders' perceptions of dimensions of care at mammographic screening programmes. It fulfils methodological standards for achieving an accurate representation of participants' perceptions. The different dimensions of satisfaction are measured and scored separately. Likert scaling allows greater flexibility in expressions of experiences than dichotomous "yes" or "no" responses, and the scales are balanced to control for potential bias from acquiescent response set.

In psychometric analysis where different dimensions of a general factor are postulated it is desirable to demonstrate statistically through factor analysis that the dimensions are differentiated. We were not able to do this because of the restricted range of responses from the test sample, where virtually all attenders were highly satisfied with all aspects of the service. We could not find and would not deliberately create the dissatisfied subset of respondents that would have been necessary to demonstrate factor analytic separation of factors.

We are reassured, however by the unanimity of the results obtained by the Q sort judges, and the results of the regression analysis that the dimensions of the scale are relevant and distinctive.

The instrument is a potentially useful tool for assessing the quality of care at mammographic screening services, and could be used routinely by such services to monitor satisfaction. As adequate attendance is critical for the success of mammography screening services, we need to ensure that women are being offered an acceptable service. By assessing participant satisfaction in a reliable and valid fashion the service can react sensitively to the needs of women. This is particularly crucial given that word of mouth about the programme is an important method by which information is spread about the service. If the hearsay is not favourable then there is the danger that potential participants will be lost. Similarly, if women are not satisfied with the service offered at initial screening then they are unlikely to attend for repeat screens.

The questionnaire has been developed specifically to look at the satisfaction with processes involved at initial screening. We have not included an assessment of the effect of outcome on perceptions of satisfaction with the screening clinic visit. For most women (well over 90% of those screened in well established programmes) no abnormality is detected, so one would not expect the small percentage of "false positives" to materially affect the overall perceptions of the initial screening process. However, it is acknowledged that a false positive result, whereby women without cancer are recalled for further tests, might have an impact on perceptions of service. In order to examine the special circumstances of this group of women, we are developing another questionnaire which measures satisfaction with the service delivery at recall assessment clinic. This questionnaire will allow important questions to be answered about the effect of false positive results on women's perceptions of a mammography screening service.

The questionnaire described in this paper is particularly useful as it allows the important dimensions of satisfaction to be examined and judged separately. By so doing, we are able to pinpoint exactly what aspects of delivery may be causing problems for participants, and consequently the areas requiring modification. Following from this, once changes in delivery are implemented, the satisfaction of participants can be reliably monitored to ensure that changes have been effective.

1. Waterhouse M, Muir C, Shanmugaratnum K, Powel J, eds. *Cancer Incidence in Five Continents*, Vol IV. IARC Scientific Publications No 42. Lyon, International Agency for Research on Cancer, 1982.

2. Tabar L, Gad A, Holmberg LH, *et al.* Reduction in mortality from breast cancer after mass screening with mammography. *Lancet* 1985, ii, 829–832.

3. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Selection, follow-up and analysis in the health insurance plan study: a randomised trial with breast cancer screening. *Nat Cancer Inst Monog* 1985, 65, 65–74.

4. Cook J, Knight D, Taylor V. *NHS Breast Screening Program. Health Education Draft Guidelines.* Oxford, Screening Publications, 1989.

5. Cockburn J, Murphy B, Schofield P, Hill D, Borland R. Development of a strategy to encourage attendance at screening mammography. *Health Education Research: Theory and Practice* (in press).

6. Irwig L, Cockburn J, Turnbull D, Simpson J, Mock P, Tattersall M. Women's perceptions of screening mammography. *Aust J Public Health* 1991, 15, 24–32.

7. Rutledge DN, Hartmann WH, Kinman PO, Windield AC. Exploration of factors affecting mammography behaviors. *Prev Med* 1988, 17, 412–422.

8. UK Trial of Early Detection of Breast Cancer Group. First results on mortality reduction of breast cancer. *Lancet* 1988, ii, 411–416.
9. Ware JE and Davies AR. Behavioral consequences of consumer dissatisfaction with medical care. *Eval Program Planning* 1983, 6, 291–297.
10. Linn MW, Linn B, Stein SR. Satisfaction with ambulatory care and compliance in older patients. *Med Care* 1982, 20, 606–614.
11. Ware JE. How to survey patient satisfaction. *Drug Intell Clin Pharmacol* 1981, 15, 892–899.
12. Ware JE, Snyder MK, Wright RW, Davies AR. Defining and measuring patient satisfaction with medical care. *Eval Program Planning* 1983, 6, 247–263.
13. Filletti GI, Firman D, Sanson-Fisher RW. Patient satisfaction with primary care consultations. *J Behav Med* 1986, 9, 389–390.
14. Ware JE. *Effects of Acquiescent Response Set on Patient Satisfaction Ratings*. Santa Monica, Rand Corporation, 1977, P.5676.
15. Cronbach LJ. *Essentials of Psychological Testing*. New York, Harper and Row, 1970.
16. Anastasi A. *Psychological Testing*. New York, Macmillan, 1976.
17. Davies AR, Ware JW. *Development of a Dental Satisfaction Questionnaire for the Health Insurance Experiment*. Santa Monica, Rand Corporation, 1982, R-2712-H115.
18. Hulka BS, Zyanski SJ. Validation of a patient satisfaction scale: theory, methods and practice. *Med Care* 1982, 20, 649–653.
19. Wolf MH, Putnam SM, James SA, Stiles WB. The medical interview satisfaction scale: development of a scale to measure patient perceptions of physician behaviour. *J Behav Med* 1978, 4, 391–401.
20. Cronbach LJ. Co-efficient alpha and the internal structure of tests. *Psychometrika* 1951, 16, 297–334.
21. Helmstadter GC. *Principles of Psychological Measurement*. New York, Appleton Century Crofts, 1964.
22. SPSSX. *Users Guide*. New York, McGraw Hill, 1983.

Long-term Survivors after Salvage High Dose Chemotherapy with Bone Marrow Rescue in Refractory Germ Cell Cancer

J.P. Droz, J.L. Pico, M. Ghosn, A. Gouyette, D. Baume, G. Piot, M. Ostronoff, C. Theodore, F. Beaujean and M. Hayat

Between April 1984 and May 1985, 17 heavily pretreated patients with relapsing or refractory germ cell tumours were treated with cisplatin 40 mg/m²/day, days 1–5; etoposide 350 mg/m²/day, days 1–5; cyclophosphamide 1600 mg/m²/day, days 2–5 and autologous bone marrow transplantation on day 8 as consolidation of conventional salvage chemotherapy. None of the 11 refractory patients and 4 of the 6 responders to prior salvage treatment are long-term survivors at 68, 72, 74 and 74 months. Mean aplasia duration was 17 days and there were 7 documented episodes of septicaemia in 17 febrile patients. 1 patient died of treatment. Among the 4 survivors, 2 patients have a sustained grade II invalidating neuropathy. We conclude that this regimen is not recommended as salvage therapy in refractory patients but may be a useful consolidation treatment in patients responding to conventional salvage chemotherapy.
Eur J Cancer, Vol. 27, No. 7, pp. 831–835, 1991

INTRODUCTION

THE PROGNOSIS of advanced stage non-seminomatous germ cell tumours has been greatly improved by the introduction of

cisplatin in chemotherapeutic regimens. Nevertheless in approximatively 20% [1] of the whole population and 40–50% of poor risk patients [1, 2] a complete remission cannot be obtained. Simultaneously 10% of those who achieve a primary complete remission relapse [3]. Several salvage conventional chemotherapy regimens have been previously reported with objective responses in up to 40% [4, 5].

In order to achieve long-term complete remission in relapsed and refractory patients we developed a high dose chemotherapy regimen followed by autologous bone marrow transplantation (ABMT) as part of their salvage treatment. Preliminary studies

Correspondence to J.P. Droz.
J.P. Droz, J.L. Pico, M. Ghosn, D. Baume, G. Piot, M. Ostronoff, C. Theodore and M. Hayat are at the Department of Medicine, and A. Gouyette is at the Department of Pharmacology, Institut Gustave-Roussy, 94800 Villejuif; and F. Beaujean is at the Centre de Transfusion Sanguine, Hôpital Henri-Mondor, Créteil, France.
Revised 2 Jan. 1991; accepted 9 Jan. 1991.