BMJ Open

International phase 1 study protocol to develop a health state classification system for a preference-based measure for women with breast cancer: the BREAST-Q Utility module

Manraj Kaur ,[1] Andrea L Pusic,[2] Stefan J Cano,[3] Feng Xie,[4] Louise Bordeleau,[5] Toni Zhong,[6] Anne Klassen [7]

To cite: Kaur M, Pusic AL, Cano SJ, *et al*. International phase 1 study protocol to develop a health state classification system for a preference-based measure for women with breast cancer: the BREAST-Q Utility module. *BMJ Open* 2020;**10**:e034451. doi:10.1136/bmjopen-2019-034451

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-034451).

Received 24 September 2019
Revised 04 December 2019
Accepted 12 December 2019

For numbered affiliations see end of article.

Correspondence to
Manraj Kaur;
kaurmn@mcmaster.ca

ABSTRACT

Introduction Concerns unique to women with breast cancer can include impact of cancer on body image, sexual well-being and changes in breast appearance and sensation. These important issues are not captured by the existing generic preference-based measures (PBMs) and no breast cancer-specific PBM currently exists. This Phase 1 protocol describes a mixed-methods study to develop and validate the descriptive health state classification system for a breast cancer-specific PBM, called the BREAST-Q Utility module.

Methods and analysis A heterogeneous sample of women aged 18 years and older diagnosed with breast cancer who are undergoing or have had treatment for breast cancer will be invited to participate in qualitative interviews. Participants will be asked to describe impact of their diagnosis and treatment(s) on their health-related quality of life (HRQOL). Interviews will be audio recorded, transcribed verbatim and coded using a line-by-line approach. At the end of each interview, based on each participant's cancer treatment history, patients will complete the mastectomy, breast-conserving therapy or reconstruction module of BREAST-Q, with modified 5-point Likert scale to measure importance of the BREAST-Q concepts. Both sources of data will be analysed to identify the most important HRQOL concerns.

A conceptual framework and item pool will be developed from the qualitative dataset. Preliminary version of the BREAST-Q Utility module will be created and refined at an in-person meeting of multidisciplinary experts. Content validity of the Utility module will be examined (cognitive debriefing, expert feedback). Psychometric properties of Utility module will be evaluated in a large sample of women with breast cancer.

Ethics and dissemination The study has been approved by Hamilton Integrated Research Ethics Board, Canada. Results of this study will be presented at international conferences and published in peer-reviewed journals.

Strengths and limitations of this study

- The BREAST-Q Utility module will be the first, rigorously developed and validated breast cancer-specific preference-based measure.
- Phase 1 involves input from a large, international sample of women with breast cancer and multidisciplinary experts, which will ensure that the Utility module measures concerns important and relevant to women with breast cancer across stage (stages 0–4) and treatment (surgical and non-surgical).
- The BREAST-Q Utility module will facilitate clinical and cost-effectiveness studies of breast cancer interventions and programmes.

BACKGROUND

Annual spending on cancer care in the USA has exceeded US$125 billion and is expected to increase exponentially.[1] Breast cancer was responsible for the largest share of cancer-related spending (13%) in 2010.[1] Past data demonstrate that the rate of growth in spending for breast cancer has exceeded that observed for lung, colorectal or prostate cancer.[2 3] Breast cancer is the most commonly diagnosed malignancy in women worldwide and the leading cause of cancer-related deaths among women.[4] The survival rate for breast cancer varies by stage of breast cancer and treatments received. Non-invasive (stage 0) and early-stage invasive breast cancer (stages I and II) have higher survival rates than later stage cancers (stages III and IV).[5] For early-stage breast cancer, the median survival can be many years, if not decades. As survival increases, healthcare resource consumption and costs associated with breast cancer can accrue years after diagnosis.

Cost-effectiveness analyses (CEAs) are used to identify the optimal allocation of healthcare resources and set funding priorities.[6–9] In CEA, the costs and outcomes of a new intervention (eg, diagnostic or interventional, surgical or non-surgical) are compared with

the costs and outcomes of an alternate, usually standard, intervention for the same health condition.[6 8] The incremental outcome of the new health intervention in CEA is described in terms of gains in quantity (ie, life expectancy) or quality of life.[10] The measure that combines these attributes (ie, quality and quantity of life) into a single index is called quality-adjusted life year (QALY). A QALY is the most commonly used metric in CEA and is defined as the value of living 1 year in full or perfect health.[9 10] Several approaches exist for estimating the 'Q' (ie, health-related quality of life (HRQOL)) in the QALY, namely, rating scales, time trade-off, standard gamble or generic preference-based measures (PBMs), such as the EuroQol-5-dimension, Short Form-6-dimension or Health Utilities Index Mark 3. The use of generic PBMs is recommended by health agencies in Australia (Pharmaceutical Benefits Advisory Committee),[11] Canada (Canadian Agency for Drugs and Technologies in Health),[12] UK for England and Wales (National Institute for Health and Clinical Excellence),[13] Scotland (Scottish Medicines Consortium)[14] and other countries in Asia, Africa, Europe and Latin America.[15] Generic PBMs are intended to be applicable to all interventions and patient groups allowing for intrapopulation and interpopulation comparisons. However, for conditions such as breast cancer, most existing generic PBMs fail to capture the unique concerns of patients, such as body image and sexual well-being.

A systematic review of studies of breast cancer interventions published between 2005 and 2017 identified no breast cancer-specific PBM.[16] Our programme of research addresses this gap by developing a condition-specific PBM for breast cancer. The development of this PBM will occur in two consecutive phases: phase 1—development and validation of a breast cancer-specific health state classification system (HSCS) and phase 2—valuation survey and modelling to produce values for health states described by the HSCS. An overview of the components of each phase is shown in figure 1. This protocol describes the phase 1 mixed-methods study to develop and validate the HSCS for the breast cancer-specific PBM. This breast cancer-specific PBM will form a new module of the BREAST-Q (hereby referred to as the BREAST-Q Utility module).

METHODS AND ANALYSIS

The first phase of developing a PBM instrument is to develop a descriptive HSCS (also called 'measurement system' or 'descriptive system'). An HSCS consists of several dimensions (or attributes), where each dimension refers to an aspect of HRQOL (eg, appearance, physical symptoms, social function).[17] In a PBM, the number of dimensions is typically limited to 7±2, with each dimension usually measured by one item. The limited number of dimensions in a PBM makes it amenable to valuation using methods such as standard gamble, time-trade off or discrete choice experiments.[17] The valuation exercise (phase 2) is used to develop the preference weights that are needed for generating health utilities.

There are two main approaches to developing a HSCS: the top-down approach, where existing literature, instruments and surveys are used to generate an item pool which is then reduced by classical test theory or item response theory; and bottom-up approach, where qualitative methods are used to identify dimensions based on

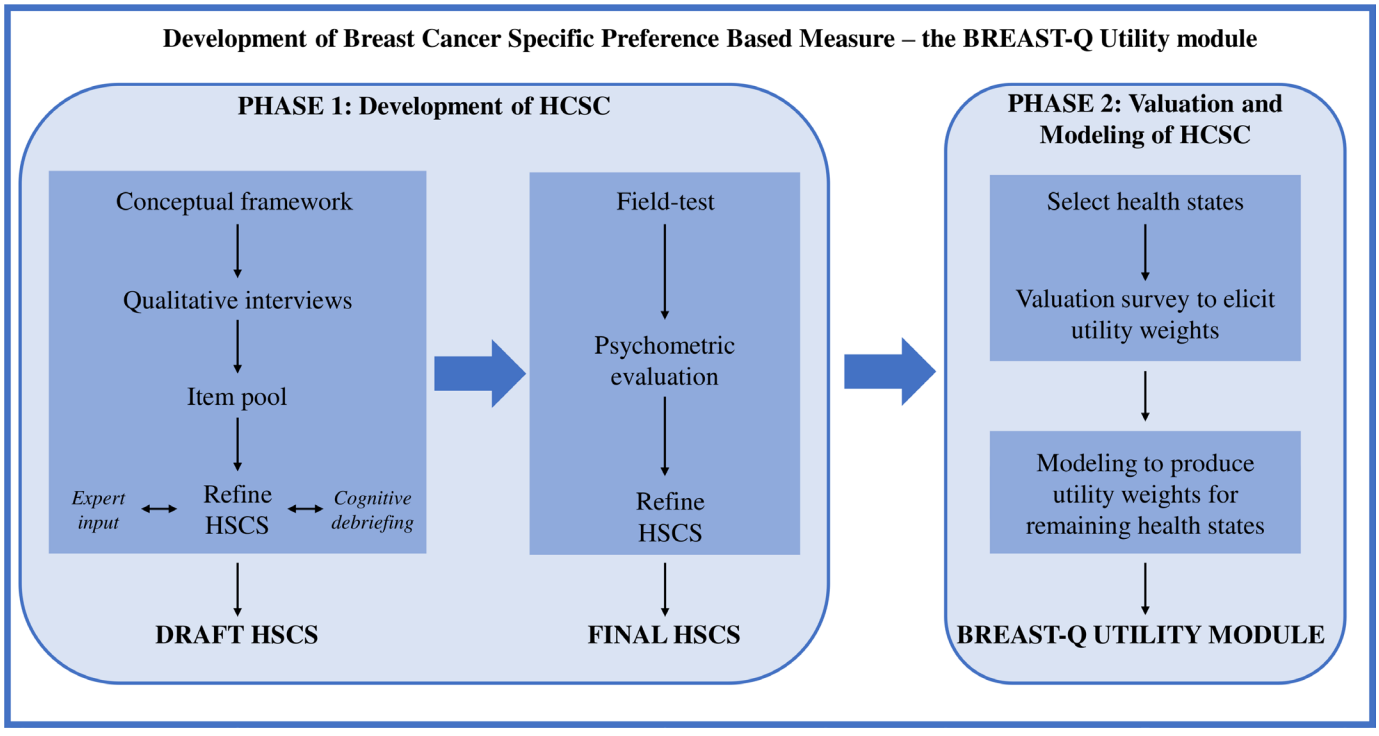

Figure 1 An overview of the multiphase, mixed-methods approach used in the development of the BREAST-Q Utility module. HSCS, health state classification system.

Figure 2 Conceptual framework of the BREAST-Q.[20]

patient input. The bottom-up approach will be used in this study. This approach is endorsed by the USA Food and Drug Administration in the development of patient-reported outcome measures (PROMs) as patient perspective is considered to be of paramount importance.[18]

An interpretive descriptive qualitative study will be conducted. Interpretive description is an inductive, analytical approach that assumes prior clinical knowledge of a health event or phenomenon being studied.[19] This approach allows for an in-depth and systematic description of a health event or phenomenon to be explored in the context of clinical knowledge and expertise to inform and guide future practice and research.[19 20]

Establishing the construct being measured

We will start with the conceptual framework of patient satisfaction and HRQOL in breast surgery that informed the development of the BREAST-Q.[21] The BREAST-Q is a PROM comprised of independently functioning scales that measure outcomes and patient's experience of care (ie, satisfaction with the surgeon, information and medical team).[21] The BREAST-Q has become the gold-standard measure of HRQOL for breast surgery and has three breast cancer surgery modules—mastectomy, breast-conserving therapy and reconstruction. The BREAST-Q conceptual framework, shown in figure 2, consists of two overarching domains: HRQOL and satisfaction with outcome. The HRQOL domain consists of three subdomains: physical, psychosocial and sexual well-being. The satisfaction with outcome domain also consists of three subdomains: satisfaction with breasts, satisfaction with overall outcome and satisfaction with care. This framework was developed from patient interviews (n=48) and refined from patient input obtained in focus groups (n=58), cognitive debriefing interviews (n=30) and from expert feedback (n=17). As such, the BREAST-Q is grounded in the patient's voice and experiences.

Health technology assessments and health policy decisions are focused on the health benefit gained from an intervention. Therefore, for a PBM to be used in QALY calculations and subsequently in CEAs, it should be able to describe and assess the health or HRQOL gain from an intervention(s). Hence, the BREAST-Q Utility module will be designed to measure the impact of breast cancer or its intervention(s) on the HRQOL of women with a diagnosis of breast cancer.

Generating an item pool

Participants, setting and recruitment

A heterogeneous sample of women with a diagnosis of breast cancer aged 18 years or older will be recruited from two breast cancer centres in Ontario, Canada (Juravinski Cancer Centre, Hamilton and Toronto General Hospital (TGH), Toronto) and one in the USA (Memorial Sloan Kettering Cancer Center (MSKCC), New York). Women undergoing diagnostic or prophylactic interventions for breast cancer or women who are unable to participate due to language barriers, cognitive or neurological deficit will be excluded from this study.

Potential study participants will be approached to participate in the study at the hospital during their clinical visit or via telephone call by a member of the clinical team within their circle of care. Patients will be provided with the study information sheet (in-person or email). After the patient has had time to review the information sheet and ask study-specific questions, their contact information will be shared with a member of the research team. Potential participants will be contacted to describe the study further and to set up a time and preferred place for the interview. Participants will be asked to sign a consent form and provide verbal consent at the start of the interview.

Sampling

We will aim to recruit a maximum variation sample[22] of women who vary by age, pathological stage of breast cancer (stages 0–4) and type and stage of surgical (ie, mastectomy, breast-conserving therapy or reconstruction) or non-surgical (ie, adjuvant or neoadjuvant therapy) breast cancer treatment. Recruitment will continue until the investigators determine that sufficient data to understand the experiences of women diagnosed with breast cancer has been obtained. According to Thorne *et al*, in interpretive description, data saturation is not the desired outcome and can be 'problematic'.[19] This is because theoretically the experiences of patients can represent an infinite number of variations.[19] Hence, the focus in interpretive description is on obtaining a deeper and richer understanding of the patient's perspective while recognizing that outliers may exist. For this study, we established a stopping criterion for data saturation as the point at which redundancy is achieved in the domains at the level of minor themes (ie, no new information is obtained). This approach is in line with the PROM development methodology.

Data collection

The BREAST-Q conceptual framework[21] will be used to develop the interview guide (box 1). Semistructured interviews[23] will be conducted in-person either at participants' homes, at the hospital in a private space or over the telephone. During the interview, we will explore

Box 1 Semistructured guide for qualitative interviews in hase 1

Experience of care
- Can you tell me about the events leading up to and including your breast cancer diagnosis?
- Can you tell me what being a breast cancer survivor/breast cancer patient has been like for you?
- What kinds of treatments have you had/are currently on/will have in the future?

Appearance
- How would you describe the appearance of your breast(s)/breast area of the chest? (Probe: with clothes, with/without bra, symmetry, contour).
- How did breast cancer and/or its treatment change the appearance of your breast(s)/breast area of the chest? (Probe: scarring, colour).
- Is there anything about your breast(s)/nipple(s)/breast area of the chest that you would like to change? (Probe: size, location, shape).
- Of the changes in your appearance, can you tell me what changes have had the most impact on you and why?

Physical function
- Do you experience any difficulty in your daily activities as a result of breast cancer/its treatment? (Probe: driving, self-care, dressing, transfers, toileting).
- Do you have any trouble moving your arm as a result of breast cancer/its treatments? (Probe: reaching objects, lifting heavy objects).
- Do you have any trouble with fine movements involving your fingers or toes? (Probe: cooking, threading a needle, typing).
- Of the physical function limitations, can you tell me which ones have had the most impact on you and why?

Physical symptoms
- Do you experience any symptoms related to breast cancer/its treatments? (eg, pain, tightness, numbness, heaviness, fatigue).
- Of the symptoms you mentioned, can you tell me which ones have had the most impact on you and why?

Psychological function
- How does breast cancer diagnosis/treatments make you feel? (Probe: upset, angry, depressed, anxious).
- How does the appearance of your breast(s)/breast area of the chest make you feel about yourself? (Probe: self-conscious, less attractive).
- Out of the emotions you mentioned, can you tell me which ones have had the most impact on you and why?

Social function
- How has breast cancer/its treatment impacted your participation in social roles? (Probe: parent, spouse, work, recreation/leisure, sports).
- Out of the social concerns you mentioned, can you tell me which ones have had the most impact on you and why?

Sexual function
- How has your sex life changed after breast cancer/its treatments? (Probe: satisfaction, feeling sexually attractive).
- Do you try to cover/hide your breasts during sex?
- Out of the sexual concerns you mentioned, can you tell me which ones have had the most impact on you and why?

Others
- Are there any other concerns or issues you experienced that we have not already covered? (eg, spiritual, coping and so on).

Most/least important aspects of health-related quality of life
- Thinking back over what you have talked about in this interview, what would you say are the top five most important aspects of your quality of life impacted by breast cancer and/or its treatment.

After completion of the BREAST-Q
- Thinking about the content of the BREAST-Q you just completed, can you tell me if there are any questions that are currently missing from the BREAST-Q that you feel are important.

the HRQOL issues most important and relevant to the participant's experience of breast cancer. Probes will be used to elicit detailed information where appropriate (eg, appearance, body image, sexual well-being). New concepts that arise will be added to the guide as the interview progress. The choice of location of the interview will depend on a patient's preference and study logistics. This will ensure that the inclusion in the study is not limited due to accessibility.

The semistructured interviews will be conducted by two experienced qualitative interviewers across the three sites. Each interview will be audio recorded and transcribed verbatim by a professional transcriptionist and identifying information will be removed. At the end of the interview, participants will be asked to describe five most important HRQOL concerns. The interviews are anticipated to last 60–90 min.

After the semistructured interview, participants will complete the most appropriate BREAST-Q module based on their surgical treatment, where the response options for the BREAST-Q scales will be replaced with a 5-point Likert scale to measure the importance of the items to the participant's experience of breast cancer (not important, slightly important, moderately important, important and very important). As the goal of our study is to develop a PBM, the BREAST-Q experience of care scales will not be completed. Participants will also be asked to nominate any items (ie, concepts) important to them that are missing from the BREAST-Q. Finally, non-identifying demographics (age, body mass index, racial or ethnic group, education level, annual income) and clinical (stage of breast cancer, type of treatments received/planned) information will be collected.

Analysis

We will use a combination of inductive (new codes developed from the data) and deductive (application of existing codes from the BREAST-Q conceptual framework)[24] to code the data. Each interview will be coded within Microsoft word using a line-by-line coding approach.[25 26] The participant quote alongside the codes will be transferred to a Microsoft Excel spreadsheet. We will also include specific participant (eg, age, country),

clinical and treatment characteristics in the Microsoft Excel spreadsheet. Constant comparison of codes will be used to refine and finalise the codes, that is, codes that have common elements will be merged to form minor themes (eg, codes about intensity, frequency, type, location and impact of pain be coded in a 'pain' category). The minor themes with common elements will be combined to form major themes (eg, pain, swelling and bruising will be grouped under 'symptoms'). The related major themes will be combined to form the top-level domain (eg, symptoms and physical function will be grouped under 'physical well-being' domain). The interview guide and codebook will be revised throughout the study as new concepts emerge. Regular team meetings will be held to review changes to the codebook. The item pool developed from the codes will be analysed to identify concepts of importance across patient, clinical and treatment characteristics. The quantitative data on the BREAST-Q item ratings of importance will be analysed descriptively using SPSS, V.25.0 (IBM Corporation for Windows/Apple Mac). Subgroup analyses using analysis of variance tests or equivalent non-parametric tests will be conducted to explore if the differences in item ratings differ by patient demographics (eg, age), clinical (eg, stage of breast cancer) or treatment (eg, type of treatment) characteristics.

The item pool will be used to draft the HSCS that contains concepts that are common across surgical and non-surgical breast cancer treatments. We will retain the language used by the participants in the wording of the items and response options. We will ensure that the item and response options are worded clearly, are easy to understand, relevant and appropriate to grade 6 reading level. Double-barreled, negatively worded or vague quantifiers will be avoided. Decisions regarding the type (eg, frequency or severity) and the number of response options to include will be guided by how the concepts are described in the qualitative data.

Credibility

To enhance credibility, several techniques will be used as follows: (a) use of audio recording and verbatim transcription by a professional transcriptionist: this will ensure errors in transcription; (b) pilot coding: the first 10 interviews (or as many as necessary) will be coded independently by two members of the research team who have experience in qualitative data analysis. The two coders will meet to review their codes, establish consensus on the definition of codes and to create a codebook. Once consistency in coding is achieved, the remaining interviews will be coded by one team member; (c) ongoing feedback: the transcribed interviews will be reviewed by a senior team member (AK) who will provide feedback on maintaining or improving quality of data collection by improving questions, altering probes or providing strategies to pursue specific aspects in greater detail; (d) member checking: the concepts elicited during the interviews will be confirmed in subsequent interviews by the interviewer; (e) debriefing: the results of the data analysis will be discussed with team members routinely via teleconference and triangulation: the conceptual framework, qualitative and quantitative data and review of the literature will be used to develop the HSCS for the Utility module.

Determining the format for measurement and response options

Once the interviews are analysed and saturation is determined to be reached, an international group of multidisciplinary experts will be invited to a 1 day, in-person meeting to review the sample characteristics, codes, item pool and draft the Utility module that covers key aspects of the preliminary conceptual framework. Feedback on attributes to be included in the Utility module and suggestions for scale items and response options will be obtained. The wording of the items and response options and the ordering of the items of the existing generic PBMs used in breast cancer research will be also reviewed.[16]

Refining the preliminary scale

A draft of the Utility module will be shown to patients and experts knowledgeable in the content area. This step will ensure that the content validity of the scale is maximised.

Cognitive debriefing interviews: patients

Participants who took part in qualitative interviews and consented to ongoing participation in the study will be invited to participate in cognitive interviews. Feedback will be obtained on the module's instructions, items and response options using the 'Think Aloud' approach.[27 28] In the think-aloud approach, participants are asked to complete each item and describe their thinking process behind choosing their response. Participants will also be asked to describe the item in their own words, what the item (ie, attribute) means to them and to provide examples from their daily activities pertinent to the item. Feedback will be obtained on the clarity and readability of the overall instrument and participants will be asked to nominate items that are missing from the Utility module. The cognitive interviews (anticipated to last 60 min) will be conducted by two experienced qualitative interviewers over phone or in-person. Interviews will be audio recorded and transcribed verbatim.

We will use the line-by-line coding approach to extract data relevant to the instructions, items and response options. The participant quote and the feedback will be transferred to Microsoft Excel worksheet for analysis. The feedback will be used to refine the instructions and response options and to decide whether to keep, modify or delete each item.[29–31] Two or three rounds of cognitive interviews will be conducted with 5–15 participants per round. Changes will be made to the Utility module after each round. The endpoint of the cognitive interviews will be when three consecutive patients do not recommend any new changes to the items in the Utility module.

Expert input

Once cognitive interviews are completed, a group of international multidisciplinary experts in the field of HRQOL and/or breast cancer research who are known to the investigators (medical and radiation oncologists, oncoplastic surgeons, allied health professionals, health economics and outcomes researchers and patient advocates) will be invited to review the Utility module using Research Electronic Data Capture (REDCap),[32] a secure web-based data collection system. Feedback on the scale instructions, items and response options will be sought and experts will be invited to nominate items that should be added or removed. Feedback will be summarised descriptively and used to make changes to the module.

PRETESTING BREAST-Q UTILITY MODULE

The BREAST-Q Utility module will be completed by a large sample of women with breast cancer (stages 0–4, any treatment). Items will be analysed in relation to demographic and clinical variables to identify the best subset of items to include in the final set of items.

Participants and recruitment

We will use the Avon Army of Women (AOW) registry to recruit women (18 years or older) who have been diagnosed with breast cancer and are fluent in English. Women undergoing prophylactic treatments for breast cancer or who have language barriers or cognitive impairments that limit participation in the study will be excluded.

Data collection

All research participants on the AOW registry will be sent an e-blast with the link to the study information sheet. Women who agree to participate will be directed to a REDCap survey to complete the BREAST-Q Utility module and a set of comparison measures (table 1). Demographic and clinical information (eg, stage of breast cancer, type of treatments to date/planned) will be collected. Participants who consent to ongoing participation will be invited to complete the BREAST-Q Utility module 1 week later to assess test–retest reliability. This time interval is sufficiently long to minimise recall bias and sufficiently short to reduce the possibility of change in responses as a result of the participant's health condition.[33] Patients will be asked if their health status is 'better', 'the same' or 'worse' since the initial administration of questionnaires.

Data analysis

The data from REDCap will be exported to SPSS, V.25.0 (IBM Corporation for Windows/Apple Mac) for analysis. The demographic and clinical characteristics of the participants will be analysed descriptively—mean (with SD) or medians (with IQR) will be used for continuous variables and percentages and frequencies will be used for categorical variables.

Psychometric evaluation of the Utility module will be performed according to the COnsensus-based Standards for the selection of health Measurement INstruments guidelines.[33] We will evaluate reliability (test–retest reliability) and construct (hypothesis testing and known groups) validity. We will also evaluate distribution of responses by items, stage of cancer and type of treatment and floor and ceiling effects (≥15% of the responses on either end of the scale[33]). Further, we will consider missing items (and reasons for missing data), descriptive feedback from participants and clinical considerations to finalise the descriptive

Table 1 Comparison measures used in the psychometric evaluation of the BREAST-Q Utility module

Measure	Characteristics
EQ-5D-5L[34–36]	► Generic preference-based measure. ► Consists of a descriptive system and the EQ-Visual Analogue Scale (VAS). The descriptive system comprises of five HRQOL dimensions with five levels each—mobility, self-care, usual activities, pain/discomfort and anxiety/depression. ► Most common utility instrument used in breast cancer research[16] Health State Utility Values in Breast Cancer: A Systematic Review of Literature).
EORTC-QLQ-C30[37]	► Cancer-specific HRQOL instrument that consists of nine multi-item scales—five functional scales (physical, role, cognitive, emotional and social); three symptom scales (fatigue, pain and nausea and vomiting); and a global health and quality-of-life scale. ► Used to derive EORTC-8D,[38] a preference-based single index measure that consists of eight dimensions—physical functioning, role functioning, pain, emotional functioning, social functioning, nausea, fatigue, sleep disturbance and constipation and diarrhoea, with four levels each (except physical functioning which has five levels).
SF-12[39]	► Generic HRQOL instrument that consists of 12 questions and 8 domains—pain, mental health, physical functioning, social functioning, role limitations due to physical and emotional problems, vitality and general health. ► Used to derive SF-6D,[40] generic preference-based measure that comprises of six domains—pain, mental health, physical functioning, social functioning, role limitations and vitality, with 4–6 response levels each.

EQ-5D-5L, EuroQol-5 dimension-5 level; HRQOL, health-related quality of life; SF-12, Short Form 12.

Table 2 Psychometric tests and criteria used in the evaluation of the BREAST-Q Utility module

Psychometric property	A priori hypothesis	Tests and criteria
Reliability The extent to which a measurement is consistent and free from error		
Test–retest reliability—the degree to which repeated measurements in stable individuals (ie, no clinical/life change) provides similar answers.[33] Measurement error—the systematic and random error of a patient's score that is not due to true changes in the construct to be measured.[33]	The BREAST-Q Utility module will demonstrate high test–retest reliability, that is, the responses between the first and second administration (1 week later) will be similar.	Weighted kappa ≥0.70[33 41] Percentage of positive and negative agreement.
Construct validity The degree to which scores of an instrument are consistent with the hypotheses, if the new instrument validly measures the construct of interest		
Hypothesis testing—the degree to which the scores of an item/scale are consistent with a priori hypothesis.[33]	Direction and magnitude of the correlation between BREAST-Q Utility module and the comparison instruments—We hypothesise that ► The BREAST-Q Utility module score will show positive (≥0.3) correlation with similar domains on EQ-5D-5L, EORTC-QLQ-C30 and SF-12. Known groups validity—Based on published evidence on HRQOL outcomes in breast cancer,[42–45] we hypothesise that the BREAST-Q Utility module score will be: ► Higher (ie, worse HRQOL) in women currently undergoing (neo)adjuvant treatment(s) compared with women who have not had/ had neoadjuvant treatment(s) in the past for breast cancer. ► Lower for women who are had breast cancer surgery alone as compared with women who had breast cancer surgery and (neo)adjuvant treatments. ► Lower for women diagnosed with early versus advanced stage breast cancer.	ANOVA or Kruskal-Wallis depending on the distribution of the data for differences in mean scores ($p<0.05$). Pearson's r or Spearman's r depending on the distribution of the data: ≥0.5 will be considered strong correlation, 0.3–0.49, moderate and 0.10–0.29 small.[33 46 47]
Acceptability and data quality		
Response distributions of the instruments and missing data Floor and ceiling effects: >15% of[33] respondents scoring the lowest or highest possible score.	We hypothesise that the Utility module will have less than 15% missing data. We hypothesise that the responses of the Utility module will be evenly distributed across the response categories (ie, no floor or ceiling effect).	Distribution of responses by instrument, item-level, stage of cancer and type of treatment will be summarised using descriptive statistics (mean, SD, % of item-level missing data).

ANOVA, analysis of variance; EQ-5D-5L, EuroQol-5 dimension-5 level; HRQOL, health-related quality of life; SF-12, Short Form 12.

HSCS of the BREAST-Q Utility module. Table 2 describes the psychometric tests and criteria that will be used in the evaluation of the BREAST-Q Utility module.

Final cognitive debriefing (post-field-test): patients

A new set of cognitive interviews will be conducted and participants will be shown the refined version of the Utility module based on the field-test results. Feedback will be obtained on the final set of items. The procedure outlined in the cognitive debriefing section will be repeated and the Utility module will be refined and finalised.

Limitations

A limitation of our study is that the interviews and field test study will involve women who are fluent in English and live in Canada or the USA. Another limitation is that participants will be drawn from a small number of cancer centres. Consequently, the results of our study may not be generalisable to women diagnosed with breast cancer in non-English-speaking countries (mainly middle-income and low-income countries). Future research will be needed to translate the Breast-Q Utility module for use in different contexts and languages.

SUBSEQUENT PHASE: VALUATION SURVEY AND MODELING TO PRODUCE VALUES FOR HEALTH STATES

Once the descriptive health state classification system of the BREAST-Q Utility module is finalised, utility weights for the health states will be developed using established methods such as standard gamble, time-trade off or

discrete choice experiments. The design of the valuation study will be determined once the health state classification system of the Utility module is validated.

PATIENT AND PUBLIC INVOLVEMENT

Our patient-centred approach engages women with breast cancer and healthcare providers in all stages of our research as experts and research team members. The use of qualitative methods ensures that the issues most important to women with breast cancer are included in the BREAST-Q Utility module. Ongoing engagement of patients in this research is ensured by inviting women who participated in the initial interviews to take part in scale refinement interviews to ensure that the Utility module is easy is understand, relevant and comprehendible. Furthermore, healthcare providers will be involved in the stages of protocol development, recruitment, data analysis and dissemination of study findings.

ETHICS AND DISSEMINATION

The patients will be invited to participate by a member of the clinical team, but the consent will be obtained by the research coordinator to ensure there is no coercion to participation. Participation in this research is voluntary. As no intervention will be provided in the course of the study, there is no direct risk to participants. However, talking about experiences with breast cancer can evoke negative feelings and unwanted recollections. If a participant feels distressed or is determined to be at risk to self or others postinterview, they will be put in touch with a skilled therapist. Participants will be made aware that they do not need to answer any question(s) that make them uncomfortable and can choose to end the interview or withdraw from the study at any time. There is no direct benefit to the participant for participating in the study except for the opportunity to contribute to improving treatment outcomes in breast cancer research. Participants in the interviews and cognitive interviews will be given a US$50 gift card as a thank you for their time.

Participants will be informed of the steps taken to protect their identity and maintain confidentiality. Any written document (eg, notes, interview transcripts, demographic forms and questionnaires) will be de-identified to ensure confidentiality. Electronic data will be stored in secure, password-protected servers and hardcopy files will be stored in a locked cabinet at the senior researcher's office at McMaster University, Canada.

The results of this study will be published in a series of articles in peer-reviewed scientific journals and presented at local, national and international conferences or meetings. Once developed, the BREAST-Q Utility module will be made available free-of-cost to non-profit users (eg, clinicians, researchers and students). Information on use, scoring and interpretation of BREAST-Q Utility module will be posted on the Q-portfolio webpage (www.qportfolio.org).

Author affiliations
[1]School of Rehabilitation Science, McMaster University, Hamilton, Ontario, Canada
[2]Plastic Surgery, Brigham and Women's Hospital, Boston, Massachusetts, USA
[3]Modus Outcomes, Letchworth Garden City, UK
[4]Department of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Ontario, Canada
[5]Medical Oncology, Juravinski Cancer Centre, Hamilton, Ontario, Canada
[6]Plastic Surgery, Toronto General Hospital, Toronto, Ontario, Canada
[7]Pediatrics, McMaster University, Hamilton, Ontario, Canada

Twitter Manraj Kaur @ThisIsManraj and Anne Klassen @anneklassen

Acknowledgements We would like to thank Ayse Kuspinar, Trisia Brietkopf and Charlene Rae for their comments and suggestions to improve the manuscript.

Contributors MK, AK, SJC and ALP were responsible for conceptualising the project, designing the study and drafting of the manuscript and approved the version that is submitted. FX, LB, SJC and TZ were responsible for designing the study and providing intellectual feedback on the manuscript. All listed authors approved the manuscript to be submitted.

Funding Phase 1 of this study was supported by the Canadian Breast Cancer Foundation Project Grant (now integrated into Canadian Cancer Society)—grant number 319371. The article processing charge was paid from this grant. MK was supported by Canadian Institutes of Health Research Canada's Best Graduate Scholarship Doctoral Award (2015-19).

Competing interests This new BREAST-Q scale will be jointly owned by Memorial Sloan-Kettering Cancer Center and McMaster University. ALP, AK and SJC are codevelopers of other BREAST-Q modules and receive a share of licence revenues based on the inventor sharing policies. The other authors have no competing interests to report in relation to the content of this article.

Patient consent for publication Not required.

Ethics approval This study is approved by the Hamilton Integrated Research Ethics Board, Hamilton, Canada. Ethics approval for the study (Project no. 2078) and data sharing agreement is also obtained from the ethics review boards and legal department of TGH (Project no. 16–5934) and MSKCC (Project no. 17-147A(1)), respectively. The ethical aspects of this research will comply with the guidelines of the national granting councils (The Canadian Institutes of Health Research) as outlined in the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Manraj Kaur http://orcid.org/0000-0002-1911-0395
Anne Klassen http://orcid.org/0000-0003-4720-0096

REFERENCES

1 Mariotto AB, Yabroff KR, Shao Y, *et al*. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst* 2011;103:117–28.
2 Warren JL, Yabroff KR, Meekins A, *et al*. Evaluation of trends in the cost of initial cancer treatment. *J Natl Cancer Inst* 2008;100:888–97.
3 Yabroff KR, Lund J, Kepka D, *et al*. Economic burden of cancer in the United States: estimates, projections and future research. *Cancer Epidemiology Biomarkers Prev* 2011;20:2006–14.
4 Servick K. Breast cancer: a world of differences. American association for the advancement of science 2014.
5 Breast Cancer – Statistics.. (2019, August 22). Retrieved October 18, 2019, from. Available: https://www.cancer.net/cancer-types/breast-cancer/statistics
6 Gold MR, Siegel JE, Russell LB. *Cost-Effectiveness in health and medicine*. Oxford university press, 1996.
7 Musgrove P, Fox-Rushby J. Cost-Effectiveness analysis for priority setting. disease control priorities in developing countries. 2006;2.

8 Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977;296:716–21.
9 Russell LB*et al*. The role of cost-effectiveness analysis in health and medicine. *JAMA* 1996;276:1172–7.
10 Pinto-Prades JL, Attema A, Sánchez-Martínez FI. Measuring health utility in economics. Oxford research encyclopedia of economics and finance 2019.
11 PBAC. *Pharmaceutical benefits Advisory Committee). guidelines for preparing submissions to the pharmaceutical benefits Advisory Committee*. Australia: Australian Government Department of Health, 2013.
12 CADTH. *Canadian agency for drugs and technologies in health). guidelines for the economic evaluation of health technologies*. Canada: CADTH, 2006.
13 NICE (National Institute of Health and Care Excellence). *Guide to the methods of technology appraisal*. London: NICE, 2013.
14 SMC (Scottish Medicines Consortium). *Guidance to manufacturers for completion of new product assessment form (NPAF)*. Scotland: NHS Scotland, 2016.
15 Rowen D, Azzabi Zouraq I, Chevrou-Severac H, *et al*. International regulations and recommendations for utility data for health technology assessment. *Pharmacoeconomics* 2017;35:11–19.
16 Kaur M. Health state utility values in breast cancer: a systematic review of literature 2019.
17 Brazier J, Ratcliffe J, Saloman J, *et al*. *Measuring and valuing health benefits for economic evaluation*. OXFORD university press, 2017.
18 U.S Department of Health and Human Services Food and Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. Available: http://www. fda. gov/ downloads/ Drugs/ Guidances/UCM193282. pdf
19 Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. *Res Nurs Health* 1997;20:169–77.
20 Thorne S, Kirkham SR, O'Flynn-Magee K. The analytic challenge in interpretive description. *Int J Qual Methods* 2004;3:1–11.
21 Pusic AL, Klassen AF, Scott AM, *et al*. Development of a new patient-reported outcome measure for breast surgery: the BREAST-Q. *Plast Reconstr Surg* 2009;124:345–53.
22 Patton MQ. *Qualitative research and evaluation methods*. 3rd ed. Thousand Oaks, CA: Sage, 2002.
23 Opdenakker R. Advantages and disadvantages of four interview techniques in qualitative research. In: *Forum qualitative Sozialforschung/Forum*. 7th vol. Qualitative Social Research, 2006: 4.
24 Crabtree BF, Miller WL, eds. *Doing qualitative research*. Sage Publications, 1999.
25 Gibbs GR. Thematic coding and categorizing. *Analyzing qualitative data* 2007;703:38–56.
26 Charmaz K, Belgrave LL. Grounded theory. The Blackwell encyclopedia of sociology 2007.
27 Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res* 2003;12:229–38.
28 Van Someren MW, Barnard YF, Sandberg JA. *The think-aloud method: a practical approach to modeling cognitive*. London: Academic Press, 1994.
29 Terwee CB, Prinsen CAC, Chiarotto A, *et al*. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27:1159–70.
30 Patrick DL, Burke LB, Gwaltney CJ, *et al*. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value Health* 2011;14:978–88.
31 Patrick DL, Burke LB, Gwaltney CJ, *et al*. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value Health* 2011;14:967–77.
32 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
33 Terwee CB, Bot SDM, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
34 Rabin R, Charro Fde, Fd C. EQ-SD: a measure of health status from the EuroQol group. *Ann Med* 2001;33:337–43.
35 EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
36 Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
37 Aaronson NK, Ahmedzai S, Bergman B, *et al*. The European organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–76.
38 Rowen D, Brazier J, Young T, *et al*. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value in Health* 2011;14:721–31.
39 Ware Jr JE, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 1996;1:220–33.
40 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851–9.
41 Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–20.
42 Pandey M, Singh SP, Behere PB, *et al*. Quality of life in patients with early and advanced carcinoma of the breast. *Eur J Surg Oncol* 2000;26:20–4.
43 Hamer J, McDonald R, Zhang L, *et al*. Quality of life (QOL) and symptom burden (SB) in patients with breast cancer. *Support Care Cancer* 2017;25:409–19.
44 McClelland SI, Holland KJ, Griggs JJ. Quality of life and metastatic breast cancer: the role of body image, disease site and time since diagnosis. *Qual Life Res* 2015;24:2939–43.
45 Jagsi R, Li Y, Morrow M, *et al*. Patient-Reported quality of life and satisfaction with cosmetic outcomes after breast conservation and mastectomy with and without reconstruction: results of a survey of breast cancer survivors. *Ann Surg* 2015;261:1198–206.
46 Nunnally JC. Psychometric theory 3E. TATA McGraw-Hill education 1994.
47 Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.