

# BC-PROM: validation of a patient-reported outcomes measure for patients with breast cancer

Xiaojuan Hu, MD, Chunsen Zhang, MD, Yanbo Zhang, PhD\*

## Abstract

Key issues of treatment for patients with breast cancer such as patient satisfaction with treatments, compliance, and some side effects are essential for evaluating quality of life by patient-reported outcomes in clinical trials. The study aimed to develop and evaluate a specific patient-reported measure which included physical, psychological, social, and therapeutic domain for assessing the survival of patients with breast cancer.

The pool of items was drafted after a theoretical revision and cognitive interviews with women with breast cancer. The draft scale was formed after the adjustment of the items and dimensions, and the selected items were submitted to expert's judgments. Five statistical methods were used to select these items by 2 validation samples. The final scale was administered to a sample of 417 patients from 8 hospitals and 135 controls for reliability, validity, and responsive analyses.

The final BC-PROM consisted of 52 items, 13 subdomains, and 4 domains, being developed after preexamination and formal examination. Cronbach alpha coefficient was 0.902 and 0.712 for the full scale and therapeutic domain. The structural validity results showed that the multidimensional measurement of the scale fulfilled expectations. Differences in the BC-PROM mean scores were significant between cancer patients and healthy participants in 13 subdomains ( $P < .05$ ), indicating good responsiveness. Among the sample survey of patients, the scale copy acceptance rate was 98.2%, completion rate 94.6%, and average filling time 10 minutes.

The new and reliable BC-PROM was developed in patients with breast cancer and applied to clinical treatment evaluation and clinical trials for such patients.

**Abbreviations:** BC-PROM = patient-reported outcomes measure for breast cancer, CFA = confirmatory factor analysis, CTT = classical test theory, ICC = item characteristic curve, IRT = item response theory, PRO = patient-reported outcome, PROM = patient-reported outcomes measure, QOL = quality of life.

**Keywords:** breast cancer, instrument development, item response theory, patient-reported outcomes, psychometrics

## 1. Introduction

Breast cancer is one of the most important female cancers in many countries, and its incidence has increased in recent decades.<sup>[1]</sup> That rise is attributed to the more widespread use of mammography and a consequent earlier detection of the disease. Earlier detection—and therefore earlier treatment—is expected to result in improved survival rates.<sup>[2]</sup> The improvement in survival rates has been related to the concern with quality of life (QOL) among the surviving breast cancer patients.<sup>[3]</sup> With advanced treatment, the role of the patient shifts to that of survivor, and there is a need for continued focus on overall QOL issues.<sup>[4]</sup>

Generally, the evaluating indices of QOL mainly include 2 aspects: information from physician and patient-rated indices. There is some difference between physician and patient-rated indices in survival status of patient with breast cancer. The patient rating is viewed as more useful because it includes subjective information only patients can provide. There is a paradigm shift in evaluating the outcomes of medical care in the past years. The outcome assessment has a more focus on the patients' perception of their health than clinical indexes of disease activity. The term "patient-reported outcome" (PRO) has been used to denote the inclusion of the patient's viewpoint in medical care.<sup>[5]</sup> The US Food and Drug Administration (FDA) 2006 draft guidance "Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims" has engendered wide discussion about PRO domains that should be taken as the end points in clinical trials.<sup>[6]</sup> PRO refers to outcomes that arise directly from patients' perceptions of their own health conditions.<sup>[7]</sup> In the case of cancer, PRO is used to assess the overall burden of the disease and the effectiveness of interventions.

There has been increased the use of outcome measures in both clinical practice and research to capture data regarding a patient's self-reported level of disability and determine the relative effectiveness of interventions.<sup>[8]</sup> The European Organization for Research and Treatment of Cancer-Breast Module and Functional Assessment of Cancer Therapy-Breast (FACT-B) were widely translated and used by many countries to measure QOL in breast cancer population.<sup>[9]</sup>

Due to the number of items and domains, scoring, and psychometric properties, there are great differences in these

Editor: Chunxiao Li.

Funding/support: This work was supported by the National Natural Science Foundation of China [81273180].

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

From the Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi, China.

\* Correspondence: Yanbo Zhang, Department of Health Statistics, School of Public Health, Shanxi Medical University, 56 South XinJian Road, Taiyuan, Shanxi Province 030001, China (e-mail: sxmuzyb@126.com).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2017) 96:17(e6781)

Received: 9 November 2016 / Received in final form: 25 March 2017 /

Accepted: 7 April 2017

<http://dx.doi.org/10.1097/MD.0000000000006781>

instruments. These items of QOL are more important in breast cancer survivors. Apart from some clinical symptoms, patients with breast cancer experience a series of nonclinical symptoms. Feelings of depression and isolation are mostly experienced. Additionally, patients with breast cancer are so fatigued that they have difficulty in daily life. The research findings have indicated that such factors as fatigue, anxiety, body image, and sexuality strongly affect QOL following breast cancer diagnosis and treatment.<sup>[10]</sup> Developments in therapeutic interventions and humanistic concerns have transformed the prospects for breast cancer patients. PRO for a patient with a history of breast cancer assess the impacts of disease, treatment, and side effects related to different treatment modalities on various aspects of a patient's outcomes.<sup>[11]</sup> The interventions of providing social support could improve QOL.<sup>[9]</sup>

Such patient-reported outcomes measure for breast cancer (BC-PROM) includes the following: measures of physical functioning, emotional functioning, social functioning, and therapeutic aspects.<sup>[12]</sup> The most indexes of patient-reported outcomes measure (PROM) are subjective indicators, and strongly influenced by the regional economic and cultural background in the evaluation of health status. Under the circumstance, there is no recognized and accurate scale of patients with breast cancer. So development and validation of a new specific BC-PROM which suit regional economic and cultural background is indispensable. The information collected from the scale is used to evaluate the effect of cancer therapy, and screening the best choices of treatment and new anticancer drugs in clinical oncology.

## 2. Methods

### 2.1. Ethics statement

The study protocol (No.2013098) received medical and ethical approval from the Shanxi Medical University. We obtained written informed consent from all participants.

### 2.2. Study population

We enrolled patients from 8 different hospitals in Shanxi Province, China. The investigators had studied training manuals and acquired investigative techniques through unified training. All participants were requested to complete the scale independently after receiving a brief introduction about the BC-PROM from one of the investigators. If a patient was unable to complete the questionnaires by themselves, they received assistance from an individual who had a proper understanding of the patient's condition. If the patients or their assistants encountered any problems with the questions, investigators were able to provide a detailed explanation. The questionnaires were filled in so as to reflect the patient's current situation. If the patients or their assistants were illiterate or had a low educational level, an investigator read out the questions and options, though without providing any actual guidance. Depending on the patient's own wishes, they could receive the assistance of an investigator or an assistant in completing the questionnaire.

Prior to the study, the investigators received relevant medical knowledge about general surgery and oncology. Furthermore, the authors made every effort to ensure that the investigators adopted a serious, responsible attitude so as to maintain the quality of this study. To minimize any missing data, the investigators checked the questionnaires immediately after completion.

The inclusion criteria for the breast cancer patients were as follows: having a definite diagnosis of breast cancer; being over 18 years of age; and being willing to undergo the investigation. The exclusion criteria were as follows: a patient with mental illness or disturbance of consciousness; inability to understand or complete the scale owing to deficiencies of language or cognitive ability; and being unwilling to undertake the investigation. Control participants were collected to meet the following criteria: not suffering from breast cancer, cancer, and mental illness; the overall age of healthy people was similar with that of patients with breast cancer; and being a volunteer to join in the research of this subject. Health controls also provided informed consent and got some rewards.

We tested the missing data in the questionnaires using Little missing completely at random test, and set a  $P$  value of  $<.001$ . If the data were being missing at random, we replaced them based on the expectation-maximization algorithm.<sup>[13]</sup>

### 2.3. Sample size

In order to obtain stable and reliable analysis results and accurate parameter estimates, some scholars have suggested that the actual survey data sample should be 5 to 10 times of the observed variables. Nunnally<sup>[14]</sup> suggested that the number of subjects was at least 10 times that of the study variables in factor analysis. For the first item-selection process, we recruited 149 breast cancer patients; valid data were obtained from 137 participants. We selected 102 patients and 35 controls, who participated in the presurvey. For the 2nd item-selection process and validation of the BC-PROM, we recruited 446 breast cancer patients and 141 controls from the same 8 regions. Of those, 417 patients and 135 controls were able to complete the final scale.

### 2.4. Scale scoring

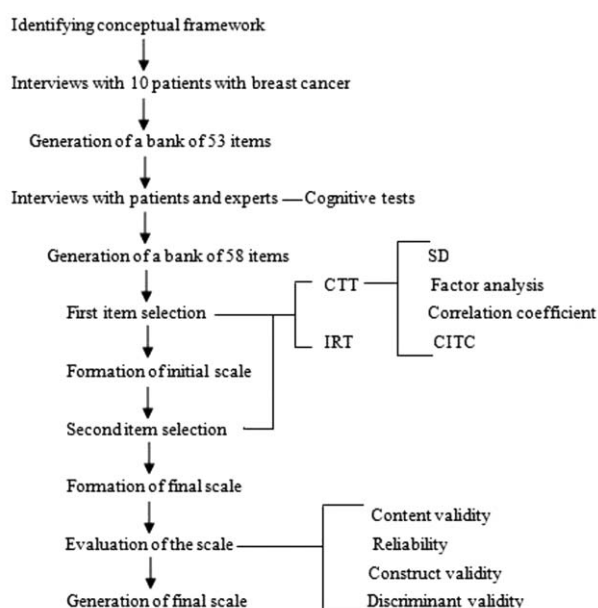
For each item, patients responded using a 5-point Likert scale to reflect how often they experienced the issues in past 2 weeks. We assigned initial values to each category, ranging from 0 to 4. The responses were 0=never, 1=occasionally, 2=about half of the time, 3=often, and 4=almost every day. Scores of positively worded items were recoded as the original score plus 1; scores of negatively worded items were recoded as 5 minus the original response. This recoding produced a score range for each item of 1 to 5, with a higher score reflecting a more positive PROM.

### 2.5. Development and formation of BC-PROM

The PROM for breast cancer was developed in 4 phases: conceptual framework construction and preliminary item generation; development of the initial scale by the 1st item-selection process; formation of the final scale based on the 2nd item-selection process; and validation of the PRO measure. A flowchart of this 4-phase developmental process appears in Fig. 1. We mainly used the methods of classical test theory (CTT) and item response theory (IRT) in the 2 item-selection processes.

### 2.6. Conceptual framework construction

Following the principles for PRO measurement tools provided by the FDA, we conducted a literature search of academic databases for PRO instruments and available Net resources for breast cancer. We then formed a conceptual framework for the new instrument. We developed 4 domains and 13 subdomains as



**Figure 1.** Flowchart of the patient-reported outcomes measures for breast cancer developmental process. CIRC= corrected item-total correlation, CTT= classical test theory, IRT= item response theory, SD= standard deviation.

follows: physical domain (subdomains: breast symptom, chest symptom, systemic symptom, unique reaction, and independence), psychological domain (subdomains: anxiety and depression, self-abasement, and despair), social domain (subdomains: social support and social adaptation), and therapeutic domain (subdomains: satisfaction, compliance, and drug side effects).

## 2.7. Item generation

We conducted in-depth open-ended interviews of 10 breast cancer patients to identify potential items for the BC-PROM using the selected conceptual framework.<sup>[15]</sup> Patients were interviewed to discuss their main physical feelings and symptoms, psychological and social burden, and satisfaction with treatment. We recorded the main points of information from those interviews.

We selected 10 patients to take a cognitive test. Those patients were requested to indicate items that they found vague or difficult to understand. Based on the patients' suggestions, we added or removed some items. As a result, we generated a bank of some potential items. We made revisions to the scale following interviews with these experts, and the preliminary scale was formed.

## 2.8. Statistical analysis

**2.8.1. Item selection.** We used the methods of CTT and IRT to evaluate the items. An item was considered for selection when it was retained by 4 or more methods. However, the evaluation also included a consideration of each item's practical significance. Together, the statistical results and practical significance of the items contributed to improvements of the scale.<sup>[16]</sup>

**2.8.2. Classical test theory (CTT).** We used 4 methods to evaluate the items in the CTT analyses. We assessed the sensitivity of items using the standard deviation (SD): an item was deleted if its standard deviation was  $\leq 1$ . We employed factor analysis to assess the subdomains. Items with low factor loading ( $< 0.40$ ) and

with factor loading close to other factors were deleted. An item was considered for deletion when the Pearson correlation between the item and its superior factor was  $< 0.50$ . Internal consistency was assessed with Cronbach alpha coefficient and the corrected item-total correlation. An item was considered for deletion when the corrected item-total correlation was  $< 0.50$  and the item's deletion increased the Cronbach alpha coefficient.<sup>[17]</sup>

**2.8.3. Item response theory (IRT).** IRT models have been the preferred methodology for statistically analyzing survey responses and patients' latent traits.<sup>[18]</sup> We assessed items using MultiLog 7.03 with Samejima graded response model to investigate the measurement properties. The graded response model is suitable for analyzing ordered response categories, such as Likert-type rating scales.<sup>[19]</sup> We used plots of an item's characteristic curves to demonstrate efficiency. Each item was characterized by 2 parameters, namely discrimination parameter and difficulty parameter. We determined these with maximum likelihood estimation. The practical values of the item parameters for deletion were as follows:  $a < 0.4$ ;  $b (-3, 3)$ .<sup>[20]</sup>

**2.8.4. Validation of final scale.** We evaluated the final BC-PROM for validity, reliability, and responsiveness using the data obtained from those 417 patients as well as the 135 control participants.

**2.8.5. Reliability.** We calculated Cronbach alpha coefficients for 4 domains and the total scale to measure the internal consistency of the BC-PROM. Generally, a value of more than 0.70 indicates that individual items provide an adequate contribution to the overall scale.<sup>[21]</sup>

**2.8.6. Validity.** Content validity. The relevant literatures and subject patients' opinions were typically consulted in validating content validity which how well these items met the empirical indexes of interest.<sup>[5]</sup>

Construct validity. We subjected the factor structure of the scale to confirmatory factor analysis (CFA). The model was contrasted for relative goodness-of-fit statistics, such as the following: root mean square error approximation (values  $< 0.08$  indicate adequate fit and values  $< 0.05$  indicate a close fit of the data to the model);<sup>[22]</sup> normed fit index (values  $\geq 0.90$ ); nonnormed fit index (values  $\geq 0.90$ ); incremental fit index (values  $\geq 0.90$ ); comparative fit index (values  $\geq 0.90$ ); and root mean square residual (values  $< 0.09$ ).<sup>[23]</sup> Using LISREL 8.70, we assessed construct validity with CFA.

Discriminant validity. We determined discriminant validity by comparing the mean scores for every subdomain of the BC-PROM among control participants and the groups of breast cancer patients. We compared the differences using analysis of variance, with the significance level set at  $P < .05$ .<sup>[21]</sup>

**2.8.7. Feasibility.** We evaluated the feasibility of the BC-PROM by examining the response rate, completion rate, and response time to completion. We considered response and return rates of less than 95% inadequate and completion times of 8 to 13 minutes acceptable.

## 3. Results

### 3.1. Participant characteristics

In the 1st item-selection phase, 102 from breast cancer patients and 35 from control participants were returned. The subjects of breast cancer were 18 to 68 years old, with an average of  $45.12 \pm$

12.04 years (see Appendix 1, Supplemental Content, <http://links.lww.com/MD/B680>, which described demographic characteristic of patients and controls in the 1st item-selection phase). In the 2nd item-selection phase, 417 breast cancer patients and 135 controls agreed to participate. The subjects of breast cancer were 20 to 73 years old, with an average age of  $47.97 \pm 10.31$  years. We calculated and compared demographic information using the Z-test for continuous variables and chi-square test for categorical variables (see Appendix 2, Supplemental Content, <http://links.lww.com/MD/B680>, which described demographic characteristic of patients and controls in the 2nd item-selection phase).

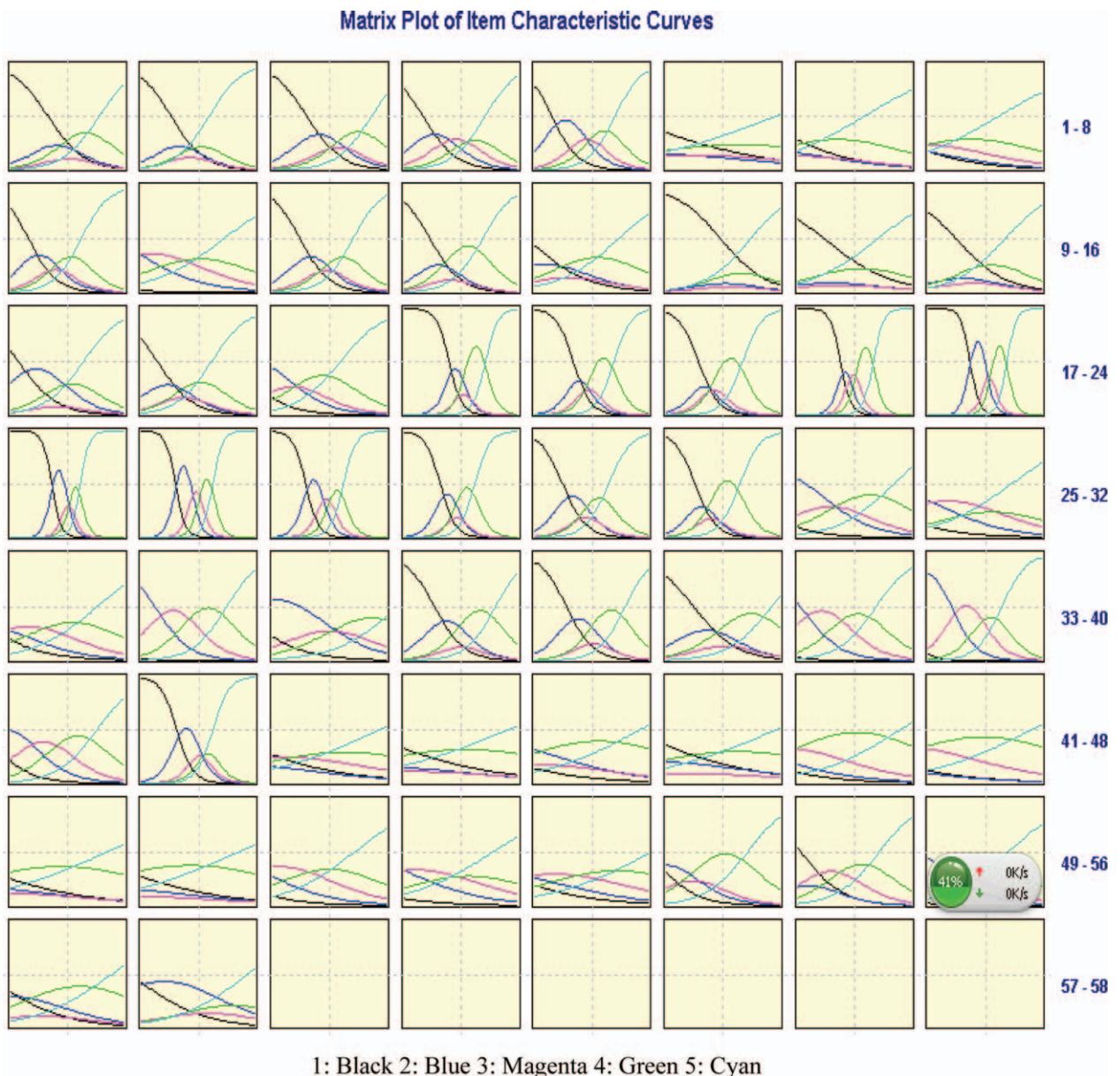
**3.2. Generation of item pool and matrix plot of item characteristic curves**

We undertook a comprehensive review to form 4 domains, 13 subdomains, and 53 potential items. Then we made an interview

with some experts to revise the various fields of the scale. We finally created a preliminary scale of 58 items. The matrix plot of item characteristic curves (ICCs) with IRT for each item is showed in Fig. 2. Each item of the BC-PROM had 5 response categories. For each item in ICCs, each response category was represented by 1 curve. Ideally, the 1st and last curve of ICCs should be monotonic for each item. And the remaining 3 curves (ie, curves of blue, magenta, and green) should be approximately normal distribution. As shown in Fig. 2, the ICCs were closer to the ideal case for a majority of items.

**3.3. Item selection**

**3.3.1. The 1st item-selection phase.** In our study, the Kaiser–Meyer–Olkin statistic was 0.706, which was larger than the gold criteria of Kaiser–Meyer–Olkin (0.7). Meanwhile the *P* value was  $<.001$  in Bartlett test of sphericity. These 2 statistics indicated the factor analysis was right for the



**Figure 2.** Item-characteristic curves and information functions using the graded-response model. The colored curves are item-characteristic curves, each corresponding to a different response category; the dashed lines are item information functions. Five curves from left to right correspond to options 1 to 5.

**Table 1**  
**Summary of the 1st item-selection phase using classical test theory and item response theory.**

Item	SD	Factor analysis	Correlation coefficient	CITC	IRT					Outcome
					a	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>	
PHD1	1.47	0.691	0.812	0.674	1.07	-1.00	-0.12	0.28	1.65	✓
PHD2	1.42	0.738	0.824	0.699	1.16	-1.43	-0.66	-0.25	0.51	✓
PHD3	1.33	0.743	0.828	0.715	1.11	-1.20	0.04	0.82	2.16	✓
PHD4	1.24	0.701	0.770	0.642	1.16	-1.99	-0.78	0.25	1.25	✓
PHD5	1.13	0.705	0.738	0.612	1.45	-2.11	-0.70	0.12	1.18	✓
PHD6	1.49	0.730	0.833	0.580	<b>0.28</b>	<b>-5.25</b>	-2.98	-0.89	2.65	✓
PHD7	1.26	0.900	0.934	0.843	0.43	<b>-5.18</b>	<b>-3.55</b>	-2.31	0.45	✓
PHD8	1.10	0.848	0.882	0.764	0.42	<b>-6.69</b>	<b>-4.51</b>	-2.21	0.62	✓
PHD9	1.14	0.791	0.854	0.766	1.37	-2.03	-0.96	-0.33	0.69	✓
PHD10	<b>0.97</b>	0.625	0.587	<b>0.432</b>	0.54	<b>-10.31</b>	<b>-3.99</b>	<b>-1.14</b>	<b>1.29</b>	×
PHD11	1.30	0.771	0.841	0.728	1.31	-1.55	-0.49	0.14	1.20	✓
PHD12	1.25	0.829	0.886	0.805	1.25	-1.61	-0.77	-0.39	1.10	✓
PHD13	1.29	0.866	0.874	0.782	0.64	-3.42	-1.69	-0.80	1.36	✓
PHD14	1.77	0.760	0.850	0.641	0.84	-0.14	0.27	0.53	1.40	✓
PHD15	1.63	0.853	0.878	0.723	0.56	-1.60	-0.96	-0.46	1.19	✓
PHD16	1.53	0.820	0.895	0.771	0.77	-1.56	-0.85	-0.39	0.99	✓
PHD17	1.30	0.793	0.898	0.745	1.00	-2.57	-0.71	-0.39	0.83	✓
PHD18	1.24	0.831	0.932	0.833	1.09	-2.17	-1.11	-0.49	0.67	✓
PHD19	0.94	0.872	0.931	0.868	0.86	<b>-4.92</b>	-2.56	-1.28	0.53	✓
PSD1	1.30	0.737	0.803	0.748	2.99	-0.63	0.00	0.23	1.28	✓
PSD2	1.19	0.568	0.700	0.630	2.20	-0.99	-0.39	0.05	1.13	✓
PSD3	1.10	0.632	0.642	0.569	1.92	-1.31	-0.74	-0.25	0.99	✓
PSD4	1.20	0.878	0.884	0.852	3.66	-0.74	-0.28	0.15	0.98	✓
PSD5	1.22	0.791	0.845	0.804	4.14	-0.83	-0.01	0.35	1.10	✓
PSD6	1.29	0.857	0.857	0.816	4.23	-0.82	-0.11	0.19	0.67	✓
PSD7	1.10	0.852	0.855	0.820	3.87	-1.22	-0.38	0.10	0.73	✓
PSD8	1.15	0.722	0.787	0.736	2.98	-1.26	-0.46	0.05	0.69	✓
PSD9	1.18	0.627	0.734	0.671	2.52	-1.08	-0.39	-0.10	0.71	✓
PSD10	1.22	<b>0.436</b>	0.611	0.524	1.65	-1.55	-0.56	-0.11	0.84	✓
PSD11	1.06	0.542	0.601	0.525	1.84	-1.49	-0.84	-0.45	0.84	✓
PSD12	1.03	0.768	0.862	0.705	0.69	<b>-6.24</b>	-2.14	-0.41	2.05	✓
PSD13	1.11	0.908	0.930	0.828	0.52	<b>-7.10</b>	<b>-3.37</b>	-0.59	1.32	✓
PSD14	1.05	0.881	0.908	0.791	0.60	<b>-5.38</b>	<b>-3.19</b>	-0.97	1.53	✓
PSD15	<b>0.84</b>	<b>0.430</b>	0.611	<b>0.477</b>	1.13	<b>-6.40</b>	-2.24	-0.46	1.45	×
PSD16	1.10	0.777	0.743	0.602	0.64	<b>-4.96</b>	-0.91	0.87	<b>3.49</b>	✓
PSD17	1.30	0.805	0.878	0.782	1.31	-1.37	-0.18	0.22	1.78	✓
PSD18	1.28	0.828	0.891	0.806	1.44	-1.33	-0.18	0.26	1.71	✓
PSD19	1.35	0.694	0.783	0.622	0.89	-1.54	-0.23	0.39	2.50	✓
SOD1	<b>0.85</b>	0.728	0.835	0.725	1.01	<b>-6.58</b>	-2.75	-0.75	1.10	✓
SOD2	<b>0.89</b>	0.710	0.896	0.817	1.54	<b>-4.84</b>	-1.77	-0.30	0.82	✓
SOD3	<b>0.94</b>	0.665	0.694	<b>0.492</b>	0.96	<b>-4.39</b>	-2.12	-0.42	1.55	×
SOD4	1.31	<b>0.612</b>	0.883	0.720	2.33	-1.15	-0.17	0.27	0.75	✓
SOD5	1.24	0.865	0.868	0.818	<b>0.33</b>	<b>-6.11</b>	<b>-4.02</b>	-1.01	2.59	✓
SOD6	1.36	0.930	0.935	0.905	<b>0.29</b>	<b>-5.53</b>	<b>-3.46</b>	-1.97	2.47	✓
SOD7	1.11	0.642	0.688	0.601	<b>0.39</b>	<b>-7.79</b>	<b>-3.58</b>	-1.80	2.51	✓
SOD8	1.44	0.696	0.755	0.655	<b>0.31</b>	<b>-5.02</b>	-2.36	-1.20	2.76	✓
SOD9	<b>0.92</b>	0.608	0.621	0.539	0.45	<b>-7.88</b>	<b>-5.18</b>	-2.18	2.41	✓
SOD10	<b>0.90</b>	0.822	0.811	0.765	<b>0.36</b>	<b>-8.61</b>	<b>-6.71</b>	<b>-2.70</b>	<b>2.41</b>	×
SOD11	1.23	0.885	0.889	0.847	<b>0.32</b>	<b>-6.41</b>	<b>-4.22</b>	<b>-2.76</b>	<b>2.05</b>	✓
SOD12	<b>0.97</b>	0.891	0.888	0.845	<b>0.32</b>	<b>-6.06</b>	<b>-4.02</b>	<b>-2.92</b>	<b>2.06</b>	×
TRD1	<b>0.97</b>	0.901	0.927	0.847	0.55	<b>-7.23</b>	<b>-4.05</b>	-1.28	1.28	✓
TRD2	1.05	0.908	0.928	0.836	0.46	<b>-8.83</b>	<b>-3.93</b>	-0.85	1.64	✓
TRD3	1.14	0.865	0.923	0.811	<b>0.39</b>	<b>-6.73</b>	<b>3.57</b>	<b>-0.44</b>	<b>3.06</b>	✓
TRD4	<b>0.97</b>	0.777	0.833	0.632	0.91	<b>-3.89</b>	-2.13	-1.08	1.25	✓
TRD5	1.09	0.823	0.899	0.730	0.91	-2.82	-1.99	-0.53	1.28	✓
TRD6	<b>0.95</b>	0.700	0.794	0.569	0.65	<b>-8.77</b>	<b>-3.23</b>	-1.18	1.17	✓
TRD7	1.27	0.674	0.830	0.431	0.51	<b>-4.33</b>	-1.88	-0.99	2.28	✓
TRD8	1.39	0.658	0.861	0.431	0.54	<b>-3.56</b>	-0.07	0.97	2.56	✓

CITC=corrected item-total correlation, IRT=item response theory, PHD=physical domain, PSD=psychological domain, SD=standard deviation, SOD=social domain, TRD=therapeutic domain. The significance of bold values in tables as per style indicated the statistical results do not meet the corresponding criteria of methods of CTT and IRT.

data. The cumulative variance contribution rate of the first 13 factors was up to 75.145%. Therefore, we selected 13 factors and assigned each item to the corresponding factors. For CTT and IRT in the 1st item-selection phase, the selection and some statistical results of the items appear in Table 1. The

items PHD10, PSD15, SOD3, SOD10, and SOD12 should have been removed in the results. However, many patients and experts agreed that there is a need to ask and investigate with regard to item SOD12 (Do you take an active part in beneficial social activities?), and so we decided to retain it. Therefore,

**Table 2**

**Summary of the 2nd item-selection phase using classical test theory and item response theory.**

Item	SD	Factor analysis	Correlation coefficient	CITC	IRT					Outcome
					a	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>	
PHD1	1.55	0.758	0.825	0.731	0.60	-2.17	-1.18	-0.43	0.94	✓
PHD2	1.54	0.861	0.884	0.819	0.54	-2.36	-1.31	-0.21	1.33	✓
PHD3	1.68	0.823	0.864	0.779	0.52	-1.15	-0.05	0.68	1.92	✓
PHD4	1.66	0.883	0.908	0.848	0.58	-1.29	-0.35	0.29	1.57	✓
PHD5	1.59	0.872	0.907	0.851	0.71	-1.48	-0.48	0.11	1.35	✓
PHD6	1.49	0.808	0.855	0.661	<b>0.38</b>	-3.73	-1.64	-0.20	2.27	✓
PHD7	1.45	0.830	0.875	0.710	<b>0.39</b>	-4.30	-1.75	-0.38	2.15	✓
PHD8	1.48	0.708	0.816	0.588	0.51	-2.16	-0.35	0.92	2.91	✓
PHD9	1.31	0.759	0.803	0.635	0.71	-2.92	-1.49	-0.29	1.43	✓
PHD10	1.20	0.733	0.835	0.705	1.02	-2.81	1.40	-0.50	0.56	✓
PHD11	1.27	0.843	0.855	0.727	0.76	-3.12	-1.53	-0.37	1.15	✓
PHD12	1.29	0.689	0.799	0.631	0.98	-2.58	-1.22	-0.32	0.50	✓
PHD13	1.33	0.681	0.817	0.589	0.53	-3.91	-1.54	-0.24	2.08	✓
PHD14	1.32	0.820	0.894	0.751	0.49	-4.50	-2.49	-1.00	1.25	✓
PHD15	1.36	0.746	0.872	0.699	0.54	-3.42	-2.49	-1.42	0.74	✓
PHD16	<b>0.99</b>	0.765	0.862	0.700	1.07	-5.50	-1.75	-0.57	1.16	✓
PHD17	1.05	0.750	0.873	0.705	0.94	-6.18	-1.58	-0.58	1.15	✓
PHD18	1.06	0.785	0.893	0.745	1.00	-4.71	-1.69	-0.76	0.70	✓
PSD1	1.36	0.734	0.757	0.692	2.35	-0.84	-0.21	0.17	1.11	✓
PSD2	1.26	0.802	0.814	0.767	2.96	-1.01	-0.24	0.17	0.99	✓
PSD3	1.24	0.917	0.908	0.884	10.27	-0.70	-0.20	0.15	0.80	✓
PSD4	1.34	0.661	0.726	0.655	1.88	0.93	-0.06	0.38	1.54	✓
PSD5	1.26	0.670	0.747	0.686	2.07	-1.26	-0.30	0.17	1.15	✓
PSD6	1.20	0.621	0.704	0.639	2.18	-1.45	-0.56	-0.05	0.81	✓
PSD7	1.20	0.685	0.753	0.696	2.66	-1.22	-0.42	0.08	0.90	✓
PSD8	1.20	0.700	0.720	0.656	2.00	-1.59	-0.35	0.19	1.20	✓
PSD9	1.22	0.904	0.887	0.858	8.48	-0.77	-0.22	0.14	0.80	✓
PSD10	<b>0.91</b>	0.577	<b>0.547</b>	<b>0.481</b>	1.48	-4.06	-1.44	-0.34	1.21	×
PSD11	1.21	0.892	0.893	0.865	9.38	-0.76	-0.23	0.15	0.81	✓
PSD12	1.57	0.772	0.885	0.702	<b>0.22</b>	-4.93	-2.74	-1.69	2.89	✓
PSD13	1.23	0.805	0.780	0.581	<b>0.37</b>	-6.59	-3.15	-1.86	2.09	✓
PSD14	1.52	0.821	0.878	0.696	<b>0.25</b>	-5.13	-2.26	-1.21	3.11	✓
PSD15	<b>0.97</b>	0.636	0.659	<b>0.463</b>	1.28	-2.50	-1.11	-0.41	0.92	×
PSD16	1.19	0.750	0.905	0.807	1.62	-1.85	-0.79	-0.08	0.90	✓
PSD17	1.14	0.717	0.834	0.687	0.77	-5.23	-2.69	-0.81	1.51	✓
PSD18	1.17	0.686	0.844	0.698	1.47	-1.90	-0.90	-0.19	1.07	✓
SOD1	1.33	0.862	0.910	0.798	0.80	-3.17	-1.01	0.06	1.41	✓
SOD2	1.28	0.873	0.921	0.826	0.81	-3.17	-1.01	0.06	1.41	✓
SOD3	1.46	0.769	0.881	0.715	0.79	-2.26	-0.25	0.37	1.41	✓
SOD4	1.05	0.812	0.807	0.729	<b>0.30</b>	-9.90	-4.05	0.25	6.77	✓
SOD5	1.09	0.764	0.795	0.709	<b>0.23</b>	-9.90	-5.52	0.93	8.77	✓
SOD6	1.02	0.625	0.618	0.492	<b>0.36</b>	-7.36	-2.21	1.77	7.69	✓
SOD7	1.10	0.750	0.759	0.660	<b>0.17</b>	14.40	6.45	0.96	12.40	✓
SOD8	1.05	0.856	0.852	0.790	<b>0.29</b>	-10.38	-4.83	0.31	6.72	✓
SOD9	1.09	0.796	0.806	0.724	<b>0.31</b>	-8.34	-3.47	1.05	6.70	✓
SOD10	1.22	0.714	0.766	0.657	<b>0.35</b>	-5.95	-2.39	0.46	4.75	✓
TRD1	1.32	0.640	0.804	0.532	0.67	-3.34	-1.97	-0.99	0.29	✓
TRD2	1.19	0.754	0.818	0.597	0.67	-4.07	-2.32	-0.97	0.74	✓
TRD3	1.29	0.677	0.800	0.534	0.80	-2.93	-1.67	-0.46	0.62	✓
TRD4	<b>0.88</b>	0.796	0.788	0.516	0.48	-7.79	-5.18	-2.15	2.17	✓
TRD5	1.02	0.766	0.818	0.502	<b>0.38</b>	-8.60	-4.52	-1.54	3.13	✓
TRD6	<b>0.77</b>	0.788	0.752	0.504	<b>0.37</b>	-11.63	-8.86	-4.04	1.91	✓
TRD7	1.09	0.878	0.898	0.592	<b>0.26</b>	-8.50	-2.62	2.17	9.31	✓
TRD8	1.04	0.883	0.887	0.592	<b>0.24</b>	-8.73	-1.86	4.06	13.34	✓

CITC=corrected item-total correlation, IRT=item response theory, PHD=physical domain, PSD=psychological domain, SD=standard deviation, SOD=social domain, TRD=therapeutic domain.

**Table 3****Scale structure of the bank of 52 items of the final scale.**

Domain	Subdomain	Item
PHD	BRS	1-, 2-, 3-, 4-, 5-
	CHS	6-, 7-, 8-
	SYS	9-, 10-, 11-, 12-,
	UNR	13-, 14-, 15-
	IND	16+, 17+, 18+
PSD	AND	1-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-
	SEA	11-, 12-, 13-
	DES	14-, 15-, 16-
SOD	SOS	1+, 2+, 3+
	SOA	4-, 5-, 6-, 7-, 8+, 9-, 10+
TRD	SAT	1+, 2+, 3+
	COM	4+, 5+, 6+
	DSE	7-, 8-

"-" indicates a reverse-scored item. AND = anxiety depression, BRS = breast symptom, CHS = chest symptom, COM = compliance, DES = despair, DSE = drug side effects, ND = independence, PHD = physical domain, PSD = psychological domain, SAT = satisfaction, SEA = self-abased, SOA = social adaptation, SOD = social domain, SOS = social support, SYS = systemic symptom, TRD = therapeutic domain, UNR = unique reaction.

we deleted 4 items from the initial scale. The scale was left with 4 domains, 13 subdomains, and 54 items.

**3.3.2. The 2nd item-selection phase.** With CTT and IRT in the 2nd item-selection phase, the selection and some statistical results of the items appear in Table 2. Items PSD10 (Have you changed your understanding and pursue of life?) and PSD15 (Are you worried about that you will die?) had to be removed according to the above criteria. Therefore, 2 items were deleted from the final scale, which consisted of 4 domains, 13 subdomains, and 52 items (see Table 3).

**3.3.3. Validation of the scale.** Additional analyses focused on examining the reliability and validity of the BC-PROM (ie, 52 items, 4 domains) using classical measurement techniques.

**3.3.4. Reliability.** Cronbach alpha coefficients for the 4 domains and overall scale appear in Table 4. As evident in that table, Cronbach alpha coefficients are 0.867, 0.884, 0.826, 0.712, and 0.902 for physical domain, psychological domain, social domain, therapeutic domain, and total scale, respectively. The 52-item BC-PROM demonstrated a high degree of internal consistency reliability.

**3.3.5. Content validity.** We achieved content validity by referring to the relevant literature. To ensure that all the items were appropriate and relevant, we consulted questionnaires from China and other countries. We also interviewed 10 patients to identify potential items; we consulted with 5 patients, 3 physician experts, and 1 psychometric expert for item revision and refinement.

**Table 4****Reliability of the 4 domains and the whole scale.**

Domain	Cronbach $\alpha$ coefficients
Physical domain	0.867
Psychological domain	0.884
Social domain	0.826
Therapeutic domain	0.712
Total	0.902

**3.3.6. Construct validity.** The indexes of fit for 4 domains are mostly presented in following. Indicators of model fitness used include the root-mean-square error approximation (0.079 for PHD, 0.13 for PSD, 0.079 for SOD, and 0.007 for TRD), normed fit index (0.94 for PHD, 0.92 for PSD, 0.96 for SOD, and 0.97 for TRD), nonnormed fit index (0.95 for PHD, 0.92 for PSD, 0.96 for SOD, and 0.98 for TRD), the comparative fit index (0.96 for PHD, 0.93 for PSD, 0.97 for SOD, and 0.99 for TRD), and incremental fit index (0.96 for PHD, 0.93 for PSD, 0.97 for SOD, and 0.99 for TRD). The fit statistics met the defined criteria, which was strongly suggested by the high factor loadings. The results of CFA appear in Table 5. The standardized factor loadings for most of the items were  $>0.5$ ; construct validity was therefore deemed satisfactory.

**3.3.7. Discriminant validity.** As shown in Table 6, the discriminant validity of each subdomain was examined by comparing mean scores in the 2nd validation samples (417 patients, 135 controls). Based on *t* tests, the rejection of the null hypothesis of each subdomain indicated that the scale had the ability to differentiate between controls and patients.

**3.3.8. Feasibility.** The response rate of the BC-PROM was 90.31%. The effective rate of return of the scale was 94.04%. The majority of participants were able to complete the scale within 10 minutes.

## 4. Discussion

With the number of women with breast cancer increasing annually among many countries, breast cancer is one of the leading causes of death in women. In such an environment, it is really essential to get more acquainted with the information of one's health-related QOL.<sup>[24]</sup> Currently, more and more patients are involved in evaluations of health-care quality using PRO.<sup>[21]</sup> The purpose of this study was to establish a reliable and valid patient-reported scale for assessing the QOL of breast cancer patients. The BC-PROM which had validated by 2 samples from 8 hospitals included areas of QOL and broader concepts, such as patient satisfaction with care and some side effects. The results of our study indicated that the BC-PROM is a valid instrument for measuring survival state for women with breast cancer.

We evaluated 4 aspects of our scale: reliability, validity, responsiveness, and feasibility. To decrease the study burden and funding, we did not evaluate test-retest reliability. We used Cronbach alpha coefficient to examine reliability. We employed the CFA method to assess construct validity. Responsiveness was evidenced by the score differences between controls and breast cancer patients. Finally, we evaluated feasibility on the basis of the effective rate of return and the time required to complete the scale. The Cronbach alpha coefficient for the breast subscale was 0.59 in Wan study on the Chinese version of FACT-B scale.<sup>[25]</sup> But the findings of this study show that the reliability is appropriate. Cronbach alpha reliability coefficient for total scale is 0.902.

During the item-selection phase, we combined CTT and IRT. CTT is easy to implement in the software of SPSS. Cronbach alpha of 4 subdomains and the whole scale was calculated which described the reliability of BC-PROM. However, CTT was item-sample dependent,<sup>[26]</sup> and its application was restricted in the development and validation of the scale. The analyses based on IRT can increase the accuracy and efficiency of the BC-PROM. Results through IRT provided much richer information on the

**Table 5****Results of the confirmatory factor analysis.**

Subdomain	Items	Nonstandardized factor loading	Standard error	t	Standardized factor loading
BRS	PHD1	1.17	0.07	17.80	0.75
	PHD2	1.28	0.06	20.61	0.83
	PHD3	1.37	0.07	20.11	0.82
	PHD4	1.50	0.06	23.51	0.90
CHS	PHD5	1.44	0.06	23.80	0.91
	PHD6	1.15	0.07	16.42	0.77
	PHD7	1.24	0.07	18.37	0.85
	PHD8	0.99	0.07	14.01	0.67
SYS	PHD9	0.93	0.06	15.72	0.72
	PHD10	0.93	0.05	17.70	0.78
	PHD11	1.02	0.06	18.43	0.80
UNR	PHD12	0.94	0.06	16.03	0.73
	PHD13	0.89	0.06	14.51	0.67
	PHD14	1.14	0.06	20.10	0.86
IND	PHD15	1.13	0.06	19.10	0.83
	PHD16	0.77	0.05	17.98	0.78
	PHD17	0.83	0.04	17.49	0.79
AND	PHD18	0.90	0.05	19.70	0.85
	PSD1	0.89	0.06	15.01	0.66
	PSD2	0.92	0.05	17.26	0.73
	PSD3	1.22	0.04	27.68	0.98
	PSD4	0.81	0.06	13.51	0.60
	PSD5	0.76	0.05	13.50	0.60
	PSD6	0.72	0.05	13.47	0.60
	PSD7	0.80	0.05	15.35	0.67
	PSD8	0.73	0.05	13.64	0.61
	PSD9	1.17	0.05	26.65	0.96
SEA	PSD10	1.17	0.04	26.71	0.96
	PSD11	1.31	0.07	17.73	0.83
	PSD12	0.79	0.06	13.51	0.65
DES	PSD13	1.24	0.07	17.32	0.81
	PSD14	1.11	0.05	23.60	0.94
	PSD15	0.88	0.05	17.99	0.78
SOS	PSD16	0.90	0.05	17.76	0.77
	SOD1	1.17	0.05	21.68	0.88
	SOD2	1.19	0.05	23.33	0.93
SOA	SOD3	1.10	0.06	17.51	0.75
	SOD4	0.81	0.04	18.15	0.77
	SOD5	0.81	0.05	17.18	0.75
	SOD6	0.54	0.05	11.06	0.53
	SOD7	0.79	0.05	16.25	0.72
	SOD8	0.89	0.04	20.94	0.85
	SOD9	0.86	0.05	18.65	0.79
	SOD10	0.85	0.05	15.69	0.70
SAT	TRD1	0.85	0.07	12.16	0.65
	TRD2	0.93	0.07	14.19	0.78
	TRD3	0.84	0.07	12.27	0.65
COM	TRD4	0.59	0.05	11.73	0.67
	TRD5	0.65	0.06	11.34	0.64
	TRD6	0.50	0.04	11.46	0.65
DSE	TRD7	1.05	0.49	2.13	0.96
	TRD8	0.64	0.30	2.11	0.62

AND = anxiety depression, BRS = breast symptom, CHS = chest symptom, COM = compliance, DES = despair, DSE = drug side effects, IND = independence, SAT = satisfaction, SEA = self-abased, SOA = social adaptation, SOS = social support, SYS = systemic symptom, UNR = unique reaction.

performance of each item.<sup>[27]</sup> It is useful during the development or refinement of our scale. ICCs provided through IRT reflected the accuracy of measurement at different values of the latent trait.<sup>[28]</sup> By examining the probability of endorsing response categories for each item, it ensured that the best items are

**Table 6****Scores comparisons between control participants and patients with breast cancer.**

Subdomain	Case (n=417)	Control (n=135)	T	P
BRS	16.73 ± 7.04	23.07 ± 3.20	10.118	<.001
CHS	9.96 ± 3.76	13.01 ± 2.78	8.698	<.001
SYS	15.40 ± 4.17	18.60 ± 2.66	8.392	<.001
UNR	11.25 ± 3.45	14.51 ± 1.02	10.841	<.001
IND	11.96 ± 2.71	13.84 ± 2.28	7.268	<.001
AND	36.33 ± 9.90	45.44 ± 6.07	10.092	<.001
SEA	10.77 ± 3.68	14.33 ± 1.37	10.980	<.001
DES	11.57 ± 3.10	14.28 ± 1.29	9.901	<.001
SOS	10.63 ± 3.68	13.19 ± 2.38	7.580	<.001
SOA	23.20 ± 5.89	29.29 ± 3.75	11.287	<.001
SAT	11.89 ± 3.07	12.56 ± 1.84	2.379	.018
COM	12.05 ± 2.10	12.68 ± 1.42	3.239	.001
DSE	5.96 ± 1.90	8.18 ± 1.33	12.557	<.001

AND = anxiety depression, BRS = breast symptom, CHS = chest symptom, COM = compliance, DES = despair, DSE = drug side effects, IND = independence, SAT = satisfaction, SEA = self-abased, SOA = social adaptation, SOS = social support, SYS = systemic symptom, UNR = unique reaction.

remained. Graded response theory was one of several models in IRT. It was used on the characteristics of ordered polytomous categories in our paper.

Many existing instruments measuring QOL have undergone development and validation among patients.<sup>[12]</sup> Referring to the pros and cons of those currently available PRO instruments, we developed a PROM to assess the survival of breast cancer patients. This study examined measurement properties, such as the reliability, validity, and responsiveness of the instrument. Finally, we developed a scale that consisted of 4 domains and 13 subdomains as follows: physiology, psychology, society, and treatment. The helpful information gathered by this instrument is used to be as prognostic and medical factors. Studies of BC-PROM can further indicate the directions needed for more efficient treatment of cancer patients. The scientific evidence from QOL data can be used to assist in clinical decision making during the phase of diagnosis and treatment for breast cancer.<sup>[29]</sup>

The BC-PROM covered more symptoms and treatment-related side effects than the Functional Assessment of Cancer Therapy–Breast Symptom Index. In the present study, the questionnaires were validated in the breast cancer population of community health centers in 3 Chinese cities. However, owing to the difficulties in following up patients and the desire not to impose an excessive burden on patients, we did not measure the test–retest reliability in the validation process of the BC-PROM. Thus, internal consistency was used only to illustrate the reliability of this scale. However, we did conduct reliability evaluation of the items with respect to both item selection and evaluation of the scale.

In 1st item-selection phase of this study, the sample size was less than 5-fold the number of selected items, and so factor analysis could not be used to explore the 4 subdomains. We need to explore this problem in future studies.

The study participants were female breast cancer patients. Because of the special symptoms of this disease, the items we assessed were related to the physiological characteristics of women. There are thus limits to the application of our scale. Owing to limited funds and other resources, our study population was restricted to Shanxi Province in northern China. Further development and wider application of the BC-PROM should be validated using nationwide or non-Chinese samples in future studies.<sup>[12]</sup>



## 5. Conclusion

This study makes an important contribution to the treatment of breast cancer and the wider health-measurement fields. We used mixed methods (CCT, IRT, and CFA) to identify beneficial items in the BC-PROM, which may be widely applied in other health areas. The validated BC-PROM could be applied to evaluate clinical treatment and clinical trials of new medicines for breast cancer patients.

### 5.1. Other Information

All authors declared that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. The manuscript had been read and approved by the authors, that the requirements for authorship as stated in the Uniform Requirements for Manuscripts Submitted to Biomedical Journals had been met.

## Acknowledgements

The authors thank the cooperation of the community hospitals at Taiyuan City, Linfen City and Jincheng City. The authors also thank the National Natural Science Foundation of China [81273180] for the support.

## References

- [1] Eskelinen M, Ollonen P. Psychosocial risk scale (PRS) for breast cancer in patients with breast disease: a prospective case-control study in Kuopio, Finland. *Anticancer Res* 2009;29:4765–70.
- [2] Thompson HS, Valdimarsdottir HB, Winkel G, et al. The Group-Based Medical Mistrust Scale: psychometric properties and association with breast cancer screening. *Prev Med* 2004;38:209–18.
- [3] Michels FA, Latorre Mdo R, Maciel Mdo S. Validity, reliability and understanding of the EORTC-C30 and EORTC-BC23, quality of life questionnaires specific for breast cancer. *Rev Bras Epidemiol* 2013;16:352–63.
- [4] Spagnola S, Zabora J, BrintzenhofeSzoc K, et al. The Satisfaction with Life Domains Scale for Breast Cancer (SLDS-BC). *Breast J* 2003;9:463–71.
- [5] Brady MJ, Cella DF, Mo F, et al. Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. *J Clin Oncol* 1997;15:974–86.
- [6] Cleland CS, Sloan JA, Group AO. Assessing the Symptoms of Cancer Using Patient-Reported Outcomes (ASCPRO): searching for standards. *J Pain Symptom Manage* 2010;39:1077–85.
- [7] Lee C, Sunu C, Pignone M. Patient-reported outcomes of breast reconstruction after mastectomy: a systematic review. *J Am Coll Surg* 2009;209:123–33.
- [8] Cook CE, Cleland J, Pietrobon R, et al. Calibration of an item pool for assessing the disability associated with foot pain: an application of item response theory to the Manchester Foot Pain and Disability Index. *Physiotherapy* 2007;93:89–95.
- [9] Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review of the literature from 1974 to 2007. *J Exp Clin Cancer Res* 2008;27:32.
- [10] Chopra I, Kamal KM. A systematic review of quality of life instruments in long-term breast cancer survivors. *Health Qual Life Outcomes* 2012;10:14.
- [11] Thomson HJ, Winters ZE, Brandberg Y, et al. The early development phases of a European Organisation for Research and Treatment of Cancer (EORTC) module to assess patient reported outcomes (PROs) in women undergoing breast reconstruction. *Eur J Cancer* 2013;49:1018–26.
- [12] Kanatas A, Velikova G, Roe B, et al. Patient-reported outcomes in breast oncology: a review of validated outcome instruments. *Tumori* 2012;98:678–88.
- [13] Preda C, Duhamel A, Picavet M, et al. Tools for statistical analysis with missing data: application to a large medical database. *Stud Health Technol Inform* 2005;116:181–6.
- [14] Nunnally JC. *Psychometric Theory*. New York: McGraw-Hill; 1967.
- [15] Baxter NN, Goodwin PJ, McLeod RS, et al. Reliability and validity of the body image after breast cancer questionnaire. *Breast J* 2006;12:221–32.
- [16] Zhi L, Qiaojun L, Yanbo Z. Development and validation of patient-reported outcomes scale for hypertension. *Int J Qual Health Care* 2015;27:369–76.
- [17] Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000;38(9 Suppl):I128–42.
- [18] Thomas ML, Brown GG, Thompson WK, et al. An application of item response theory to fMRI data: prospects and pitfalls. *Psychiatry Res* 2013;212:167–74.
- [19] Emons WH, Meijer RR, Denollet J. Negative affectivity and social inhibition in cardiovascular disease: evaluating type-D personality and its assessment using item response theory. *J Psychosom Res* 2007;63:27–39.
- [20] Abdel-Aal RE, El-Alfy E-SM. Constructing optimal educational tests using GMDH-based item ranking and selection. *Neurocomputing* 2009;72:1184–97.
- [21] Meads DM, McKenna SP, Doward LC, et al. Development and validation of the Asthma Life Impact Scale (ALIS). *Respir Med* 2010;104:633–43.
- [22] Broekman BF, Niti M, Nyunt MS, et al. Validation of a brief seven-item response bias-free geriatric depression scale. *Am J Geriatr Psychiatry* 2011;19:589–96.
- [23] Luckett T, King MT. Choosing patient-reported outcome measures for cancer clinical research—practical principles and an algorithm to assist non-specialist researchers. *Eur J Cancer* 2010;46:3149–57.
- [24] Perry S, Kowalski TL, Chang CH. Quality of life assessment in women with breast cancer: benefits, acceptability and utilization. *Health Qual Life Outcomes* 2007;5:24.
- [25] Patoo M, Allahyari AA, Moradi AR, et al. Persian Version of Functional Assessment of Cancer Therapy- Breast (FACT-B) Scale: confirmatory factor analysis and psychometric properties. *Asian Pac J Cancer Prev* 2015;16:3799–803.
- [26] Kahler CW, Strong DR, Hayaki J, et al. An item response analysis of the Alcohol Dependence Scale in treatment-seeking alcoholics. *J Stud Alcohol* 2003;64:127–36.
- [27] Hagman BT, Kuerbis AN, Morgenstern J, et al. (IRT) analysis of the Short Inventory of Problems-Alcohol and Drugs (SIP-AD) among non-treatment seeking men-who-have-sex-with-men: evidence for a shortened 10-item SIP-AD. *Addict Behav* 2009;34:948–54.
- [28] Neal DJ, Corbin WR, Fromme K. Measurement of alcohol-related consequences among high school and college students: application of item response models to the Rutgers Alcohol Problem Index. *Psychol Assess* 2006;18:402–14.
- [29] Harrington S, Michener LA, Kendig T, et al. Patient-reported upper extremity outcome measures used in breast cancer survivors: a systematic review. *Arch Phys Med Rehab* 2014;95:153–62.